Dissertation

STOCHASTIC DIFFERENTIAL EQUATIONS WITH SINGULAR DRIFTS AND MULTIPLICATIVE NOISES

Chengcheng Ling

Fakultät für Mathematik
Universität Bielefeld

September 2019

Acknowledgement

I am very grateful to my supervisor Prof. Dr. Michael Röckner from Bielefeld University. During my three years PhD program, he gave me a lot inspiring advice and continuous encouragement on my research. In addition to valuable knowledge about mathematics, patience and persistence are another two important things that I learnt from him.

I am very grateful to Prof. Dr. Zhiming Ma. He always encourages me step by step and keep going, which motivated me all the time. Under his recommendation, I got this precious opportunity to study in Bielefeld University.

I would like to thank Prof. Dr. Xiangchan Zhu, Prof. Dr. Rongchan Zhu. They patiently and meticulously guide me how to do research. With their help, I learnt a lot on the road of scientific research. I also got many opportunities to learn from excellent mathematicians. It widely broads my horizon, and it also helps me enjoy the research very much.

I would like to thank Prof. Dr. Kassmann, Prof. Dr. Kumagai, Prof. Dr. Trutnau for their support and valuable suggestions. I am very appreciated about the strong support from DFG through the IRTG 2235 Bielefeld-Seoul 'Searching for the regular in the irregular: Analysis of singular and random systems'. I would like to express my appreciation to the 6 months exchange stay in Seoul National University. Many sincerely thanks for all of the professors and Phd students and post doctor students in IRTG 2235 group. I had a very nice time staying with them and enjoyed research very much.

I would like to thank Prof. Dr. Deng Zhang, Dr. Guohuan Zhao, Dr. Longjie Xie, Dr. Xiaobing Sun, Dr. Xue Peng, Huanyu Yang and Bingguang Chen for helpful discussions. In particular I would like to thank my joint workers Dr. Guohuan Zhao and Dr. Longjie Xie, I got a lot valuable knowledge about stochastic differential equations and partial differential equations after many times of discussion with them. I am also very appreciated about the useful advice and help from Dr. Xue Peng for my writing of this thesis.

I would like to thank Prof. Dr. Fengyu Wang, Prof. Dr. Xicheng Zhang, Prof. Dr. Zhenqing Chen for their support and useful advice on research.

I also would like to thank Rebecca Reischuk, Nadine Brehme, Anke Bodzin, Claudia Köhler for their help in Bielefeld.

In the end I give my sincerely thanks to the support from my family and friends.

Best wishes for all of the people who I ever met!

Preface

This thesis mainly talks about stochastic differential equations (abbreviated as SDEs) with singular drifts and multiplicative noise. The following four aspects about SDEs are considered.

- The well-posedness of SDEs driven by continuous multiplicative noise on $\mathbb{R}_+ \times \mathbb{R}^d$ in mixed norm space. We obtain the existence and uniqueness of a strong global and continuous solution to SDE in mixed norm space.

- The well-posedness of SDEs driven by continuous multiplicative noise on a general space time domain $Q \subset \mathbb{R}_+ \times \mathbb{R}^d$ in mixed norm space. We prove the maximally defined existence and uniqueness of strong solutions to SDEs driven by multiplicative noise on general space-time domains $Q \subset \mathbb{R}_+ \times \mathbb{R}^d$, which have continuous paths on the one-point compactification $Q \cup \partial$ of Q where $\partial \notin Q$ and $Q \cup \partial$ is equipped with the Alexandrov topology.

- The non-explosion of the solutions to SDEs driven by continuous multiplicative noise obtained on general space time domains $Q \subset \mathbb{R}_+ \times \mathbb{R}^d$ in mixed norm space. We prove that under some Lyapunov type conditions, the explosion time of the solution to SDE with gradient type drift is infinite and its distribution has sub-Gaussian tails.

- The well-posedness of SDEs driven by jump processes ($\alpha-$stable like processes) with distributional valued drifts. We show the well-posedness of nonlocal elliptic equation with distributional-valued drift in Besov-Hölder spaces first. Then we obtain the existence and uniqueness for corresponding martingale problem, which is equivalent to the existence and uniqueness of weak solution to SDE. Moreover, we prove that the one dimensional distribution of the weak solution has a density in some Besov space.

Contents

1 Introduction

1.1 Background and Motivation

A stochastic differential equation (abbreviated as SDE) is used in engineering and physics to describe how random factors ('noise') can be incorporated into classical dynamical equations. We consider the following equation on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$ on $[0,\infty)\times\mathbb{R}^d$:

$$\begin{aligned} X_t = x + \int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s \\ + \int_0^t\int_{|z|<1} g(s,X_{s-},z)\tilde{N}(ds,dz) + \int_0^t\int_{|z|\geqslant 1} g(s,X_{s-},z)N(ds,dz), \quad t\geqslant 0, \end{aligned} \tag{1.1}$$

with measurable coefficients $b : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d\times\mathbb{R}^d$, $g : [0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}^d$, $(0,x)$ is the starting point, and $(W_t)_{t\geqslant 0}$ is a $d-$dimensional $(\mathcal{F}_t)-$Brownian motion defined on this probability space, and N is an $(\mathcal{F}_t)-$Poisson random measure with intensity measure $dt\nu(dz)$, where ν is a Lévy measure on $\mathbb{R}^d$, that is

$$\int_{\mathbb{R}^d} (|z|^2\wedge 1)\nu(dz) < \infty, \quad \nu(\{0\}) = 0,$$

and the compensated Poisson random measure $\tilde{N}$ is defined as

$$\tilde{N}(dt,dz) := N(dt,dz) - dt\nu(dz).$$

Usually we call $\int_0^t b(s,X_s)ds$ the drift term, $\int_0^t \sigma(s,X_s)dW_s$ the continuous noise term, and $\int_0^t\int_{\mathbb{R}^d} g(s,X_{s-},z)N(ds,dz)$ the jump type noise term of the SDE (1.1). If $\sigma\equiv 0$ and $g\equiv 0$, SDE (1.1) becomes an ordinary differential equation (abbreviated as ODE):

$$x'(t) = b(t,x(t)), \quad x(0) = x. \tag{1.2}$$

Thus, we can treat a stochastic differential equation as a generalization of an ordinary differential equation by adding the effect of noise. An interesting phenomenon of (1.1) is that the noise term plays some regularization effect such that the SDE (1.1) is well-posed for quite singular drifts b. For instance, the ODE (1.2) does not have a unique solution if b is merely Hölder continuous (say $d=1$, and $b(x) := |x|^\alpha$ for some $\alpha\in(0,1)$), but if we add a Brownian motion to (1.2), we can obtain the uniqueness of the solution almost surely in probability. In the past decades, there is an increasing interest in the study of the SDE (1.1). In this thesis, we mainly study the following aspects about SDEs, we give each of them an introduction.

1.1.1 Well-posedness of SDEs

Firstly, we want to study the well-posedness of SDEs (i.e. existence and uniqueness of the solutions to SDEs) with singular coefficients. With regard to whether the noise of the SDE is continuous or allows jumps, we divide our introduction into the following two parts.

- **SDE driven by continuous noise**

We consider the following SDEs driven by continuous noise (i.e. $g \equiv 0$ in SDE (1.1)):

$$X_t = x + \int_0^t b(s+r, X_r)dr + \int_0^t \sigma(s+r, X_r)dW_r, \quad t \geqslant 0, \tag{1.3}$$

in an open subset $Q \subset \mathbb{R}^{d+1}$.

There are many known results on studying existence and uniqueness of strong solutions to the SDE (1.3). In the seminal paper [62], Veretennikov proved that when $Q = \mathbb{R}_+ \times \mathbb{R}^d$, if the coefficient σ is Lipschitz continuous in the space variable x uniformly with respect to the time variable t, $\sigma\sigma^*$ is uniformly elliptic, and b is bounded and measurable, then the SDE (1.3) admits a unique global strong solution (i.e. $\xi = \infty$ $a.s.$ where ξ is the lifetime of the solution $(X_t)_{t\geqslant 0}$). In [37], under the assumptions that: $\sigma = \mathbb{I}_{d\times d}$ ($\mathbb{I}_{d\times d}$ denotes the unit matrix in $\mathbb{R}^d$) and $bI_{Q^n} \in L^{q(n)}(\mathbb{R}; L^{p(n)}(\mathbb{R}^d))$ for $p(n), q(n) \in (2,\infty)$ and $d/p(n) + 2/q(n) < 1$, where Q^n are open bounded subsets of Q with $\overline{Q^n} \subset Q^{n+1}$ and $Q = \cup_n Q^n$, Krylov and Röckner proved the existence of a unique maximal local strong solution to the SDE (1.3) when Q is a subset of $\mathbb{R}^{d+1}$, which says that there exists a unique strong solution $(s+t, X_t)$ solving the SDE (1.3) on $[0,\xi)$ such that $[0,\infty) \ni t \to (s+t, X_t) \in Q' := Q \cup \partial$ (Alexandrov compactification of Q) is continuous and this process is defined to be in ∂ if $t \geqslant \xi$. To this end they applied the Girsanov transformation to get existence of a weak solution firstly and then proved pathwise uniqueness of (1.3) by Zvonkin's transformation invented in [79]. Then, the well-known Yamada-Watanabe theorem [71] yields existence and uniqueness of a maximal local strong solution. Assuming that for $b \in L^q_{loc}(\mathbb{R}_+, L^p(\mathbb{R}^d))$ with $p, q \in (1,\infty)$ and $d/p + 2/q < 1$ and $\sigma = \mathbb{I}_{d\times d}$, Fedrizzi and Flandoli [21] introduced a new method to prove existence and uniqueness of a global strong solution to the SDE (1.3) by using regularizing properties of the heat equation. This method was extended by von der Lühe to the multiplicative noise case in her work [64]. Zhang in [73] proved existence and uniqueness of a strong solution to the SDE (1.3) on $Q = \mathbb{R}_+ \times \mathbb{R}^d$ for $t < \tau$, where τ is some stopping time, under the assumptions that σ is bounded, uniformly elliptic and uniformly continuous in x locally uniformly with respect to t, and $|b|, |\nabla\sigma| \in L^{q(n)}_{loc}(\mathbb{R}_+; L^{p(n)}(B_n))$ ($\nabla\sigma$ denotes the weak gradient of σ with respect to x) with $p(n), q(n) \in (2,\infty)$ satisfying $d/p(n) + 2/q(n) < 1$, where B_n is the ball in $\mathbb{R}^d$ with radius $n \in \mathbb{N}_+$ centering at zero. Zvonkin's transformation plays a crucial role in Zhang's proof. In [63], [72], [70] and references therein the well-posedness of the SDE (1.1) was also studied. The above results include the case where the coefficients of SDE (1.1) are time dependent. For the time independent case, Wang [65] and Trutnau [41] used generalized Dirichlet forms to get existence and uniqueness results of the SDE (1.3) on $Q = \mathbb{R}^d$.

However, the conditions imposed on the coefficients in the above mentioned results concerning the strong well-posedness for SDE ([37],[73],[21]) are not unified when we think an SDE in two ways: as a system (i.e. each component $(X^i_t)_{t\geqslant 0}, 1 \leqslant i \leqslant d$ of the vector $(X_t)_{t\geqslant 0} = (X^1_t, \cdots, X^d_t)_{t\geqslant 0} \in \mathbb{R}^d$ satisfies an SDE in $\mathbb{R}^1$) and as a whole SDE in

$\mathbb{R}^d$. Let us illustrate this by a simple example: consider the following SDE in $\mathbb{R}^2$:

$$\begin{cases} \mathrm{d}X_t^1 = b_1(t, X_t^1)\mathrm{d}t + \mathrm{d}W_t^1, & X_0^1 = x_1 \in \mathbb{R}, \\ \mathrm{d}X_t^2 = b_2(t, X_t^2)\mathrm{d}t + \mathrm{d}W_t^2, & X_0^2 = x_2 \in \mathbb{R}. \end{cases} \tag{1.4}$$

If we set $x := (x_1, x_2) \in \mathbb{R}^2$, $X_t := (X_t^1, X_t^2)$, $W_t := (W_t^1, W_t^2)$, and define the vector field

$$b(t,x) := \begin{pmatrix} b_1(t, x_1) \\ b_2(t, x_2) \end{pmatrix}.$$

Then SDE (1.4) can be rewritten as

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t, \quad X_0 = x \in \mathbb{R}^2, \quad t \geqslant 0. \tag{1.5}$$

According to the above mentioned criterion, we need to assume

$$b \in L^q_{loc}(\mathbb{R}_+; L^p(\mathbb{R}^2)) \quad \text{with} \quad 2/p + 2/q < 1$$

to ensure the well-posedness of SDE (1.5). This in particular means that we need

$$b_1, b_2 \in L^q_{loc}(\mathbb{R}_+; L^p(\mathbb{R}^1)) \quad \text{with} \quad 2/p + 2/q < 1. \tag{1.6}$$

On the other hand, the two-dimensional SDE (1.4) can also be viewed as single equations for $(X_t^1)_{t\geqslant 0}$ and $(X_t^2)_{t\geqslant 0}$ themselves, because $(X_t^1)_{t\geqslant 0}$ and $(X_t^2)_{t\geqslant 0}$ are not coupled in the equation. From this point of view, the SDE (1.4) can be well-posed under the weaker condition that

$$b_1, b_2 \in L^q_{loc}(\mathbb{R}_+; L^p(\mathbb{R}^1)) \quad \text{with} \quad 1/p + 2/q < 1,$$

which does not coincide with the obviously hence not optimal condition (1.6). The point is that we might have a non-uniformly in integrability of our coefficients with respect to the component of its variables which need to be taken into account to optimize our conditions. We point out that such non-uniformity will always appear when we consider multi-dimensional SDEs, and especially for the degenerate noise cases and multi-scale models involving slow and fast phase variables, see e.g. [22, 68].

Hence, the one of aims of our work is to take into account the above non-uniformity by studying SDEs with coefficients in general mixed-norm spaces. It turns out that the appropriate condition is

$$\frac{1}{p_1} + \cdots + \frac{1}{p_d} + \frac{2}{q} < 1,$$

where p_i is the integrability of each component $x_i \in \mathbb{R}$ of the drift b. The condition $\frac{1}{p_1} + \cdots + \frac{1}{p_d} + \frac{2}{q} < 1$ shows explicitly how much contribution comes from the time variable t in $\mathbb{R}_+$ and each component x_i of the space variable $x = (x_1, \cdots, x_d) \in \mathbb{R}^d$. Therefore, as a generalization of the classical Lebesgue space $L^p(\mathbb{R}^d)$, we will study the SDE (1.3) in the mixed norm space $L^{\mathbf{p}}(\mathbb{R}^d) := L^{p_d}(\mathbb{R}, L^{p_{d-1}}(\mathbb{R}, (\cdots, L^{p_1}(\mathbb{R}))))$ where $\mathbf{p} = (p_1, \cdots, p_d)$.

Besides, we are also interested in the existence and uniqueness of a maximally defined local strong solution to (1.3) on $Q \subset [0, \infty) \times \mathbb{R}^d$, especially when Q is not very regular, say $Q = \mathbb{R} \times (\mathbb{R}^d \backslash \gamma^\rho)$, where $\gamma^\rho = \{x \in \mathbb{R}^d | dist(x, \gamma) \leqslant \rho\}$, $\rho \geqslant 0$, and γ is a countable locally finite subset defined as $\gamma = \{x_k | k \in \mathbb{N}\} \subset \mathbb{R}^d$, where none of the above results mentioned can be applied, except for the one in [37]. However, [37] is restricted to the case where the diffusion term is a Brownian motion. We want to obtain the existence and uniqueness of a maximally defined local strong solution to (1.3) on $Q \subset [0, \infty) \times \mathbb{R}^d$ as well in the case where the diffusion matrix σ is not constant. In the end we show that if σ is bounded, uniformly elliptic and uniformly continuous in x, locally uniformly with respect to t, and $|bI_{Q^n}|, |\nabla \sigma I_{Q^n}| \in L^{q(n)}(\mathbb{R}_+; L^{p(n)}(\mathbb{R}^d))$ with $p(n), q(n) \in (2, \infty)$ satisfying $d/p(n) + 2/q(n) < 1$, there exists a maximally defined local strong solution $(s+t, X_t)_{t \geqslant 0}$ to (1.3) on Q such that $[0, \infty) \ni t \to (s+t, X_t) \in Q' := Q \cup \partial$ (Alexandrov compactification of Q) is continuous and this process is defined to be in ∂ if $t \geqslant \xi$.

• SDE driven by jump type noise

In recent years, SDEs on $\mathbb{R}^d$ driven by pure jump Lévy processes and irregular drifts have also a lot of attracted interest. For simplicity, we consider the following simplified form of SDE (1.1)

$$X_t = x + \int_0^t g(X_{s-}) dL_s + \int_0^t b(X_s) ds, \quad t \geqslant 0, \tag{1.7}$$

where L_t is an α-stable process in $\mathbb{R}^d$, g is a $d \times d$-matrix-valued measurable function and b is the drift, which might be very singular. In [60] Tanaka, Tsuchiya and Watanabe showed that if $(L_t)_{t \geqslant 0}$ is a symmetric $\alpha-$stable process with $\alpha \in (0, 1)$, b is time independent, bounded and $\beta-$Hölder continuous with $\beta < 1 - \alpha$, $g \equiv 1$, the SDE (1.7) may not have a unique strong solution. When $\alpha \in [1, 2)$, $g \equiv 1$ and $b \in \mathcal{C}_b^\beta(\mathbb{R}^d)$ with $\beta > 1 - \frac{\alpha}{2}$ Priola in [52] proved that there exists a unique strong solution to the SDE (1.7). Under the same condition, Haadem and Proske [31] obtained the unique strong solution by using the Malliavin calculus. Zhang [74] proved the pathwise uniqueness to the SDE (1.7) when $\alpha \in (1, 2)$, b is bounded and in some fractional Sobolev space. Recently in [5] Athreya, Butkovsky, and Mytnik obtained uniqueness and existence of strong solution to the SDE (1.7) when $d = 1$, $g \equiv 1$ and b is just in a certain class of Schwartz distributions. See also [13, 14, 12, 58, 69] for more results related to (1.1). Basicly these works showed that the SDE (1.7) has a unique strong solution under the conditions that g is bounded, uniformly nondegenerate and Lipschitz, L_t is an α-stable process, $b \in C^\beta$ (Hölder space) with $\beta > 1 - \frac{\alpha}{2}$. We can find that (1.7) is a special case of (1.1) with $\sigma \equiv 0$.

Since the results in [5] considered the additive noise case, i.e. $g \equiv 1$ with $d = 1$ only. A natural question is whether the well-posedeness still holds for the SDE (1.7) with multiplicative noise for $d \geqslant 1$. In order to answer this question, we study the existence and uniqueness of the solution to (1.7) with distribution-valued drift in some class and multiplicative noise in multiple dimensions.

1.1.2 Non-explosion of solutions to SDEs driven by continuous noise

We consider the following stochastic differential equation with continuous noise

$$X_t = x + \int_0^t b(s+r, X_r)dr + \int_0^t \sigma(s+r, X_r)dW_r, \quad t \geqslant 0, \tag{1.8}$$

in an open subset $Q \subset \mathbb{R}^{d+1}$ with measurable coefficients $b : Q \to \mathbb{R}^d$ and $\sigma : Q \to \mathbb{R}^d \times \mathbb{R}^d$. Here $(s,x) \in Q$ is the starting point, and $(W_t)_{t\geqslant 0}$ is a $d-$dimensional $(\mathcal{F}_t)-$ Brownian motion defined on some complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$. Define

$$\xi := \inf\{t \geqslant 0 : (t+s, X_t) \notin Q\}. \tag{1.9}$$

ξ is called the explosion time (lifetime) of the process $(t+s, X_t)_{t\geqslant 0}$ in the domain Q. If $\xi < \infty$ $a.s.$, we call the solution $(t+s, X_t)_{t\geqslant 0}$ a local solution. If $\xi = \infty$ $a.s.$, it is called a global solution. As we introduced already, we can show that there exists a maximal local strong solution to (1.8). Then it is very natural to ask in which case this maximal local solution is global.

When $Q = \mathbb{R}_+ \times \mathbb{R}^d$, there are several well-known results about non-explosion of the solution to the SDE (1.8). In [62, 63, 79] the assumptions that b and σ are bounded and Lipischitz continuous and $\sigma\sigma^*$ is uniformly elliptic guarantee that the solution will not blow up. Zhang in [72] obtained that under the conditions that σ is continuous, uniformly nondegenerate and $\sup_{t\in[0,T]} \|\nabla\sigma\|_{L^{2(d+1)}(B_n)} < \infty$, $|b| \leqslant C + F$, for some constants C and $F \in L^p([0,\infty) \times \mathbb{R}^d)$, $p > d$, the solution to (1.8) does not explode. Xie and Zhang in [70] proved that if σ is locally uniformly continuous in x and locally uniformly with respect to $t \in [0,\infty)$, and for some $q > d+2$, $b \in L^q_{loc}([0,\infty) \times \mathbb{R}^d)$, $\nabla\sigma \in L^q_{loc}([0,\infty) \times \mathbb{R}^d)$, and for some constants $C_1, \gamma_1 > 0$, $\alpha' \in [0,\alpha)$, $\alpha \geqslant 0$, and for all $t \geqslant 0$, $x \in \mathbb{R}^d$, $\xi \in \mathbb{R}^d$,

$$|\sigma(t,x)\xi| \geqslant |\xi| \Big(I_{\alpha>0} \exp\{-C_1(1+|x|^2)^{\alpha'}\} + I_{\alpha=0} C_1 (1+|x|^2)^{-\gamma_1} \Big)$$

and

$$\langle x, b(t,x)\rangle + \kappa(1+|x|^2)^{\alpha} |\sigma(t,x)|^2 \leqslant C_1(1+|x|^2),$$

there exists a unique global strong solution to the SDE (1.8). This non-explosion result was obtained by directly applying Itô's formula to an exponential function $\exp\{e^{-\lambda t(1+|x|^2)^{\alpha}}\}$ for some positive constant λ. The above results are about the case where the coefficients are time dependent. For the time independent case, Wang [65], Lee and Trutnau [41] used generalized Dirichlet forms to get non-explosion results.

As mentioned in [37], there are several interesting situations arising from applications, say diffusions in random media and particle systems, where the domain Q of (1.8) is not the full space $\mathbb{R} \times \mathbb{R}^d$ but a subdomain (e.g. $Q = \mathbb{R} \times (\mathbb{R}^d \backslash \gamma^{\rho})$, where $\gamma^{\rho} = \{x \in \mathbb{R}^d | dist(x,\gamma) \leqslant \rho\}$, $\rho \geqslant 0$, and γ is a locally finite subset of $\mathbb{R}^d$), where none of the results mentioned above can be applied to get global solutions, except for the one in [37]. Moreover, Krylov and Röckner in [37] did not only prove the existence and uniqueness of a maximal local strong solution of equation on Q, but also they obtained that if $b \equiv= -\nabla\phi$, where $\phi : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ is a nonnegative function, and if there exist a constant $K \in [0,\infty)$ and an integrable function h on Q^n, with Q^n as defined as above such that the following Lyapunov conditions hold in the distributional sense

$$2D_t\phi \leqslant K\phi, \quad 2D_t\phi + \Delta\phi \leqslant he^{\epsilon\phi}, \quad \epsilon \in [0,2), \tag{1.10}$$

the strong solution does not blow up, which means $\xi = \infty$ *a.s.*. Here $D_t\phi$ denotes the derivative of ϕ with respect to t. This result can be applied to diffusions in random environment and also finite interacting particle systems to show that the process does not exit from Q or goes to infinity in finite time.

However, [37] is restricted to the case where equation (1.8) is driven by additive noise, that is, the diffusion term is a Brownian motion. Our interest is about the case when σ is not only a constant-valued matrix. In this case, the key step is to find the appropriate Lyapunov conditions generalizing (1.10).

1.1.3 Density of the solutions to SDEs driven by jump noise

Recently Debussche and Fourier in [15] proved that there exists a density of the solution to the SDE

$$X_t = x + \int_0^t g(X_{s-})dL_s + \int_0^t b(X_s)ds, \quad t \geqslant 0, \tag{1.11}$$

where $(L_t)_{t\geqslant 0}$ is an $\alpha-$stable process with $\alpha \in (0,2)$, g and b are Hölder continuous functions, and the density lies in some Besov space. This work can be seen as a probabilistic approach to the theory of regularity of solutions to non-local partial differential equations. Indeed, the density of the solution of a stochastic equation satisfies a Fokker-Planck equation which is, in the jump case, non-local. There is a lot of research in this field in the PDE community. In particular, some results are available in the case of coefficients with low regularity. The typical result is that when the initial condition is continuous, the viscosity solution is immediately Hölder continuous, see Barles, Chasseigne and Imbert [8] and references therein. Concerning the techniques used to prove the existence of the density, based on Fourier transform and the Plancherel identity, there is one method introduced in [27] to prove the existence of a density for the time-marginals of many stochastic processes, which can be applied to study one-dimensional SDEs whose coefficients have low regularity. This method has been refined and generalized in [28] such that it can deal with multidimensional processes.

However, in [15], they did not show the existence and uniqueness of the solution to (1.11)which we as indicated above establish the well-posedness to (1.11) in this work. As a next step we want to study the existence and regularity of the density of the solution that we obtained for (1.11) with distributional-valued drift and multiplicative jump noise.

1.2 Main results

Firstly we obtain the existence and uniqueness of a strong global solution to SDE driven by continuous noise in mixed norm space $L^{\mathbf{p}}(\mathbb{R}^d) := L^{p_d}(\mathbb{R}, L^{p_{d-1}}(\mathbb{R}, (\cdots, L^{p_1}(\mathbb{R}))))$ where

$\mathbf{p} = (p_1, \cdots, p_d)$. More precisely, in **Theorem 3.1** (p.30) we show that if for some $p_1, \cdots, p_d, q \in (1, \infty]$ and every $T > 0$,

$$|b|, |\nabla\sigma| \in L^q([0,T]; L^{\mathbf{p}}(\mathbb{R}^d)) \quad \text{with} \quad \frac{2}{q} + \frac{1}{p_1} + \cdots + \frac{1}{p_d} < 1,$$

and for every $n \in \mathbb{N}$, σ is uniformly continuous in $x \in \mathbb{R}^d$ uniformly with respect to $t \in [0,T]$, and there exist positive constants δ_1 and δ_2 such that for all $(t,x) \in [0,T] \times \mathbb{R}^d$,

$$\delta_1 |\xi|^2 \leqslant |\sigma^*(t,x)\xi|^2 \leqslant \delta_2 |\xi|^2, \quad \forall \xi \in \mathbb{R}^d.$$

Then for any $(\mathcal{F}_t)-$stopping time τ and $x \in \mathbb{R}^d$, there exists a unique strong continuous solution $(X_t)_{t \geqslant 0}$ such that

$$P\left\{\omega : \int_0^T |b(r, X_r(\omega))| dr + \int_0^T |\sigma(r, X_r(\omega))|^2 dr < \infty, \forall T \in [0, \tau(\omega))\right\} = 1,$$

and

$$X_t = x + \int_0^t b(r, X_r) dr + \int_0^t \sigma(r, X_r) dW_r, \quad \forall t \in [0, \tau), a.s..$$

The condition $\frac{1}{p_1} + \cdots + \frac{1}{p_d} + \frac{2}{q} < 1$ shows the contributions on integrability of b and $\nabla\sigma$ with respect to time variable t in $\mathbb{R}_+$ and each component x_i of the space variable $x = (x_1, \cdots, x_d)$ in $\mathbb{R}^d$.

Based on the existence and uniqueness of a global strong solution to (1.3) that we obtained on $[0,\infty) \times \mathbb{R}^d$ in mixed-norm space, by applying a localization procedure, we get the existence and uniqueness of a maximally defined local strong solution in $Q' = Q \cup \partial$ (one-point compactification of Q) for $Q \subset [0,\infty) \times \mathbb{R}^d$. Our results **Theorem 4.1** (p.57) show that if for any $n \in \mathbb{N}$ and some $\mathbf{p}(n) = (p_1(n), \cdots, p_d(n))$, $q(n) \in (1, \infty)$ satisfying $1/p_1(n) + \cdots + 1/p_d(n) + 2/q(n) < 1$,

$$|bI_{Q^n}|, \quad |\nabla\sigma I_{Q^n}| \in L^{q(n)}([0,T]; L^{\mathbf{p}(n)}(\mathbb{R}^d))$$

and for $1 \leqslant i,j \leqslant d$, $\sigma_{ij}(t,x)$ is uniformly continuous in x uniformly with respect to t for $(t,x) \in Q^n$, and there exists a positive constant δ_n such that for all $(t,x) \in Q^n$,

$$|\sigma^*(t,x)\lambda|^2 \geqslant \delta_n |\lambda|^2, \quad \forall \lambda \in \mathbb{R}^d,$$

then for any $(s,x) \in Q$, there exists a unique continuous $Q'-$valued function $(z_t)_{t \geqslant 0} := (t, X_t)_{t \geqslant 0}$ and a $(\mathcal{F}_t)-$stopping time $\xi =: \inf\{t \geqslant 0 : z_t \notin Q\}$ such that $(X_t)_{t \geqslant 0}$ is the unique strong solution to the following SDE

$$X_t = x + \int_0^t b(s+r, X_r) dr + \int_0^t \sigma(s+r, X_r) dW_r, \quad \forall t \in [0, \xi), a.s. \tag{1.12}$$

and for any $t \geqslant 0$, $z_t = \partial$ on the set $\{\omega : t \geqslant \xi(\omega)\}$ $(a.s.)$.

As far as the non-explosion result is concerned, we have to take into account that having non-constant σ instead of $\mathbb{I}_{d\times d}$ in front of the Brownian motion in (1.12) means that we have to consider a different geometry on $\mathbb{R}^d$, and that this effects the Lyapunov function type condition which is to replace (1.10) and also the form of the equation. By comparing the underlying Kolmogrov operators of the SDEs, in **Theorem 5.2** (p.67) we get that the following type setting SDE should be considered ($(a_{ij})_{1\leqslant i,j\leqslant d} = \sigma\sigma^*$):

$$X_t = x + \int_0^t (-\sigma\sigma^*\nabla\phi)(s+r, X_r)dr + \frac{1}{2}(\sum_{j=1}^d \int_0^t \partial_j a_{ij}(s+r, X_r)dr)_{1\leqslant i\leqslant d} + \int_0^t \sigma(s+r, X_r)dW_r, \quad t \geqslant 0 \tag{1.13}$$

which is the 'geometrical' analogue to the additive noise case

$$X_t = x + \int_0^t (-\nabla\phi)(s+r, X_r)dr + W_t, \quad t \geqslant 0,$$

where ϕ is a non-negative continuous function on $[0,\infty) \times \mathbb{R}^d$. To be more specific, since the Kolmogrov operator $\mathcal{L}$ corresponding to (1.13) is given by

$$\mathcal{L} = div(\sigma\sigma^*\nabla) - \langle \sigma^*\nabla\phi, \sigma^*\nabla \rangle \,,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^d$. Recalling that $div \circ \sigma$ is the adjoint of the 'geometric' gradient $\sigma^*\nabla$ (i.e. taking into account the geometry given to $\mathbb{R}^d$ through σ). So, the Laplacian Δ in (1.10) is to be replaced by the Laplace-Beltrami operator $div(\sigma\sigma^*\nabla)(= \sum_{i,j=1}^d \partial_j(a_{ij}\partial_i))$ and the right Lyapunov type condition of non-explosion to the SDE (1.13) is that there exists a constant $K \in [0,\infty)$ and an integrable function h on Q^n, defined as above, such that the following conditions hold in the distributional sense

$$2D_t\phi \leqslant K_1\phi, \quad 2D_t\phi + \sum_{i,j=1}^d \partial_j(a_{ij}\partial_i\phi) \leqslant he^{\epsilon\phi},$$

which then indeed turns out to be the correct analogue to (1.10). This leads to some substantial changes in the proof of our non-explosion result in comparison with the one in [37].

For the SDE driven by jump type noise, in **Chapter 6** (p.95) we consider the following SDE

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}^d} \int_0^\infty z\mathbf{1}_{[0,\kappa(X_{s-},z))}(r) N^{(\alpha)}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) + \int_0^t b(X_s)\mathrm{d}s, \tag{1.14}$$

where κ is a nonnegative measurable function from $\mathbb{R}^d \times \mathbb{R}^d$ to $[0,\infty)$ and $N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s)$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+$ with intensity measure $\mathrm{d}r \, \frac{\mathrm{d}z}{|z|^{d+\alpha}} \, \mathrm{d}s$. We obtain the existence and uniqueness of a weak solution for very singular drift b which is maybe even only a Schwartz distributions. This result essentially follows from the existence and uniqueness of the solution to the following non-local partial differential equation

$$\lambda u - \mathscr{L}_\kappa^\alpha u - b \cdot \nabla u = f. \tag{1.15}$$

Here $\alpha \in (0,2)$, $b \in \mathscr{C}^\beta$(Besov-Hölder space, see Definition 6.7 below) with $\beta \in \mathbb{R}$, and

$$\mathscr{L}_\kappa^\alpha f(x) := \int_{\mathbb{R}^d} \left(f(x+z) - f(x) - \nabla f(x) \cdot z^{(\alpha)} \right) \frac{\kappa(x,z)}{|z|^{d+\alpha}} \mathrm{d}z,$$

where $z^{(\alpha)} := z\mathbf{1}_{\{|z|<1\}}\mathbf{1}_{\alpha=1} + z\mathbf{1}_{\alpha\in(1,2)}$. By applying the Littlewood-Paley theorem we obtain the existence and uniqueness of the weak solution to (1.15). Then by Zvonkin's transformation (which will be introduced in details in Chapter 2.5.1), we get the existence and uniqueness of a weak solution to SDE (1.14). Furthermore, by refining the method from [15] (which will be introduced in Chapter 2.5.2) we obtain existence and regularity estimates for the density of the weak solution to SDE (1.14) in Besov space.

1.3 Structure of this thesis

In order to make the thesis self-contained, in Chapter 2 we collect the basic concepts and some fundamental results which we will use subsequently.

As the generalization of the classical Lebesgue L^p space, we will study the SDEs with continuous noise in the mixed-norm space in Chapter 3.

Based on the results that we proved in Chapter 3 for $Q = [0,\infty) \times \mathbb{R}^d$, we apply the localization procedure to obtain the maximal local strong solution on general domains Q in Chapter 4. That is to say, we prove existence and uniqueness of maximally defined strong solutions to SDEs driven by multiplicative noise on general space-time domains Q in $\mathbb{R}_+ \times \mathbb{R}^d$ in mixed norm sapce, which have continuous paths in the one-point compactification $Q \cup \partial$ of Q where $\partial \notin Q$ and $Q \cup \partial$ is equipped with the Alexandrov topology. Besides, we give several examples for which we show well-posedness result in $Q' = Q \cup \partial$ by our result.

In Chapter 5, our aim is to extend the non-explosion results in [37] to the multiplicative noise case on a general domain Q. We also give two important applications from diffusions in random media and particle systems respectively. Both are generalizations of examples in [37, Section 9] to the case of multiplicative noise.

In Chapter 6 by applying Littlewood-Paley theory we first prove that there exists a weak solution to the equation (1.15) when b is very irregular. Then the existence and uniqueness of the weak solution (which is equivalent to the martingale solution) to SDE (1.14) with possibly distributional valued drifts in the multi-dimension follows from Zvonkin's transformation. Which extends the result of [76] to jump type noise and [5] to multiplicative noise. Based on the well-posedness results proved in the first part of Chapter 6, by a similar argument as in [15], which is refined in our work by applying

Littlewood-Paley theory, we obtain the existence and regularity of the density of the time marginals of the solutions to the SDE (1.14) with possibly distributional valued drifts.

The Appendix contains technical lemmas used in the proofs of our main results.

1.4 Outlook

As it is shown in the following chapters, all of the studied stochastic differential equations are assumed to have nondegenerate noise term, and until now it is not clear whether the existence and uniqueness of strong solution holds for the SDEs with possibly distributional-valued drifts when $d \geqslant 2$. Noting that Littlewood Paley theory plays a powerful role during dealing with partial differential equations with singular coefficients. Based on these considerations, there are four main topics which I would like to study in future work:

- the existence and uniqueness of strong solutions to the SDEs with distributional-valued drifts when $d \geqslant 2$,
- the existence and uniqueness of strong solutions to the SDEs with degenerate noise,
- the existence and uniqueness of strong solutions to the SDEs with a mixture of continuous noise (Brownian motion) and jump type noise ($\alpha-$stable process),
- the properties (e.g. non-explosion, strong Feller, ergodicity) of solutions to SDEs.

2 Preliminaries

Throughout this thesis, we use the following convention: C with or without subscripts will denote a positive constant, whose value may change from one appearance to another, and whose dependence on parameters can be traced from calculations.

2.1 Mixed-norm Lebesgue spaces

For the convenience of reading and also in order to make the thesis self complete, we first give a brief introduction about the mixed-norm Lebesgue spaces and collect the theorem which will be used later, for more details we refer to [3] and the references therein.

Let $\mathbf{p} = (p_1, ..., p_d) \in [1, \infty)^d$ be a multi-index, we denote by $L^{\mathbf{p}}(\mathbb{R}^d)$ the space of all measurable functions on $\mathbb{R}^d$ with norm

$$\|f\|_{\mathbf{p}} := \left(\int_{\mathbb{R}} \cdots \left(\int_{\mathbb{R}} |f(x_1, ..., x_d)|^{p_1} dx_1 \right)^{\frac{p_2}{p_1}} dx_2 \right)^{\frac{p_3}{p_2}} \cdots dx_d \right)^{\frac{1}{p_d}} < \infty.$$

Thus $L^{\mathbf{p}}(\mathbb{R}^d)$ is a Banach space. The order is important when taking the integrals in the expression above. If we permute the p_is, then increasing the order of p_i gives the smallest norm, while by decreasing the order gives the largest norm. If we define the conjugate exponent $\mathbf{p}' = (p'_1, \cdots, p'_d)$ to $\mathbf{p} = (p_1, \cdots, p_d)$ with $\frac{1}{p_1} + \frac{1}{p'_1} = 1$, for $i = 1, \cdots, d$, we write as $\frac{1}{\mathbf{p}} + \frac{1}{\mathbf{p}'} = 1$, then $L^{\mathbf{p}'}(\mathbb{R}^d)$ is the dual of $L^{\mathbf{p}}(\mathbb{R}^d)$ for $\mathbf{p} \in [1, \infty)^d$. Without any surprise, we have Hölder's inequality, Minkowski's inequality for integrals and dominated convergence theorem .

Lemma 2.1. *([3, Lemma 2])(Hölder's inequality). For any $\mathbf{p} \in [1, \infty]^d$, $f \in L^{\mathbf{p}}(\mathbb{R}^d)$ and $g \in L^{\mathbf{p}'}(\mathbb{R}^d)$, we have*

$$\left| \int_{\mathbb{R}^d} fg d\mathbf{x} \right| \leqslant \|f\|_{L^{\mathbf{p}}} \|g\|_{L^{\mathbf{p}'}}.$$

Lemma 2.2. *([3, Lemma 4])(Minkowski's inequality). For any $\mathbf{p} \in [1, \infty]^d$ and a measurable function $f \in L^{(\mathbf{p}, 1, \cdots, 1)}(\mathbb{R}^{d_1+d_2})$, we have*

$$\left\| \int_{\mathbb{R}^{d_2}} f(\cdot, \mathbf{y}) d\mathbf{y} \right\|_{L^{\mathbf{p}}} \leqslant \int_{\mathbb{R}^{d_2}} \|f(\cdot, \mathbf{y})\|_{L^{\mathbf{p}}} d\mathbf{y}.$$

Lemma 2.3. *([3, Theorem 2])(Dominated convergence Theorem). Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of measurable functions on $\mathbb{R}^d$. If $f_n \to f$ (a.e.) and if there is a dominating function $G \in L^{\mathbf{p}}(\mathbb{R}^d)$ such that $|f_n| \leqslant G$ (a.e.) for any n, then $\|f_n - f\|_{L^{\mathbf{p}}} \to 0$.*

Let $\mathscr{S}(\mathbb{R}^d)$ be the Schwartz space of all rapidly decreasing functions, and $\mathscr{S}'(\mathbb{R}^d)$ the dual space of $\mathscr{S}(\mathbb{R}^d)$. Then from the argument in [3] we have

Lemma 2.4. *([3, Throrem 3]) The following inclusions hold*

$$\mathscr{S}(\mathbb{R}^d) \hookrightarrow L^{\mathbf{p}} \hookrightarrow \mathscr{S}'(\mathbb{R}^d).$$

Furthermore, they are dense and continuous for $\mathbf{p} \in [1, \infty)$.

Given $f \in \mathscr{S}(\mathbb{R}^d)$, let $\mathscr{F} f = \hat{f}$ be the Fourier transform of f defined by

$$\hat{f}(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x) \mathrm{d}x.$$

We know that the *Fourier multiplier* T_m is a linear operator that acts by multiplying the Fourier transform $\hat{f}$ by a function m, and then applying the inverse Fourier transform $\mathscr{F}^{-1}$, which can be said that T_m reshapes the frequencies of f. For complex-valued function f on $\mathbb{R}^d$, we have

$$(T_m f)(x) := \mathscr{F}^{-1}(m\hat{f}) = \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi} m(\xi)\hat{f}(\xi) d\xi. \tag{2.1}$$

For $\mathbf{p} \in [1,\infty)^d$, we denote by $\mathcal{M}_{\mathbf{p}}$ the space of all bounded complex finctions m on $\mathbb{R}^d$ such that the operator $T_m f$, which is initially defined for $f \in \mathscr{S}$, can be extended to a bounded operator on $L^{\mathbf{p}}(\mathbb{R}^d)$, the norm is defined as

$$\|T\|_{\mathcal{M}_{\mathbf{p}}} := \|T\|_{\mathcal{L}(L^{\mathbf{p}}(\mathbb{R}^d))}.$$

Then $\mathcal{M}_{\mathbf{p}}$ is a closed subspace of $\mathcal{L}(L^{\mathbf{p}}(\mathbb{R}^d))$ and thus it is a Bananch space. The elements of the space $\mathcal{M}_{\mathbf{p}}$ are called $L^{\mathbf{p}}$ *Fourier multipliers*. In the following we give the Hörmander-Mihlin theorem for mixed norm spaces proved in [3].

Theorem 2.5. *([3, Theorem 7]) Let $m \in L^\infty(\mathbb{R}^d\backslash\{0\})$ be such that for some $A > 0$ and for any multi-index $|\alpha| \leqslant \lfloor \frac{d}{2} \rfloor + 1$, it satisfies on of the following conditions*

(a) Mihlin's condition

$$|\partial_\xi^\alpha m(\xi)| \leqslant A|\xi|^{-|\alpha|},$$

(b) Hörmander's condition

$$\sup_{R>0} R^{-d+2|\alpha|} \int_{R<|\xi|<2R} |\partial_\xi^\alpha m(\xi)|^2 d\xi \leqslant A^2 < \infty.$$

Then, m lies in $\mathcal{M}_{\mathbf{p}}$ for any $\mathbf{p} \in (1,\infty)$, and we have the estimate

$$\begin{aligned}\|m\|_{\mathcal{M}_{\mathbf{p}}} &\leqslant \sum_{k=1}^{d} c^k \prod_{j=0}^{k-1} \max\{p_{d-j}, (p_{d-j}-1)^{-1/p_{d-j}}\}(A + \|m\|_{L^\infty}) \\ &\leqslant c' \prod_{j=0}^{d-1} \max\{p_{d-j}, (p_{d-j}-1)^{-1/p_{d-j}}\}(A + \|m\|_{L^\infty}),\end{aligned}$$

where c and c' are constants that depend only on d.

2.2 Lévy processes and non-local pseudo-differential operators

The aim of this section is to give a brief introduction about Lévy processes and pseudo-differential operators which are related to the knowledge that we will use in the later chapters. We refer to [2] for more details.

Definition 2.6. *For a random variable X defined on probability space $(\Omega, \mathcal{F}, P)$ and taking values in $\mathbb{R}^d$ with distribution P_X its characteristic function $\phi_X : \mathbb{R}^d :\to \mathbb{C}$ is defined by*

$$\phi_X(\xi) = E(e^{i\xi \cdot X}) = \int_\Omega e^{i\xi \cdot X(\omega)} P(d\omega) = \int_{\mathbb{R}^d} e^{i\xi \cdot y} P_X(dy)$$

for each $\xi \in \mathbb{R}^d$.

We wrote down the characteristic function $\phi_\mu(\xi) = \int_{\mathbb{R}^d} e^{i\xi \cdot y} \mu(dy) = e^{\eta(\xi)}$ of distribution μ, we call the map $\eta : \mathbb{R}^d \to \mathbb{C}$ a *Lévy symbol.*

Definition 2.7. *For a random variable X taking values in $\mathbb{R}^d$, we say that X is infinitely divisible if, for all $n \in \mathbb{N}$, there exists i.i.d random variables $Y_1^{(n)}, \cdots, Y_n^{(n)}$ such that*

$$X \overset{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}.$$

Definition 2.8. *Let $X = (X_t)_{t \geqslant 0}$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, P)$. We say that X is a* Lévy process *if:*

1. *$X_0 = 0 (a.s.)$;*

2. *X has independent and stationary increments, i.e. for each $n \in \mathbb{N}$ and each $0 \leqslant t_1 < t_2 < \cdots < t_{n+1} < \infty$ the random variables $(X_{t_{j+1}} - X_{t_j})_{1 \leqslant j \leqslant n}$ are independent and each $X_{t_{j+1}} - X_{t_j}$ has the same distribution as $X_{t_{j+1} - t_j}$.*

3. *X is stochastic continuous, i.e. for all $a > 0$ and for all $s \geqslant 0$*

$$\lim_{t \to s} P(|X_t - X_s| > a) = 0.$$

If X is a Lévy process, then X_t is infinitely divisible for each $t \geqslant 0$, and we can write $\phi_{X_t}(\xi) = e^{\eta(t,\xi)}$ for each $t \geqslant 0$, $\xi \in \mathbb{R}^d$, where each $\eta(t, \cdot)$ is a Lévy symbol, and $\eta(\cdot)$ is the Lévy symbol of X_1. Before we give the the *Lévy-Khintchine formula*, which is the cornerstone for much of what follows, we introduce *Lévy measure* first:

Definition 2.9. *For a Borel measure ν defined on $\mathbb{R}^d \backslash \{0\}$ we say that it is a Lévy measure if*

$$\int_{\mathbb{R}^d \backslash \{0\}} (|y|^2 \wedge 1) \nu(dy) < \infty.$$

The result given below is usually called *Lévy-Khintchine formula* with respect to an infinitely divisible measure.

Theorem 2.10. *(Lévy-Khintchine) $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ is infinitely divisible if there exits a vector $b \in \mathbb{R}^d$, a positive definite symmetric $d \times d$ matrix A and a Lévy measure ν on $\mathbb{R}^d \backslash \{0\}$ such that for all $\xi \in \mathbb{R}^d$,*

$$\phi_\mu(\xi) = \exp\Big\{ib \cdot \xi - \frac{1}{2}\xi \cdot A\xi + \int_{\mathbb{R}^d\backslash\{0\}} [e^{i\xi\cdot y} - 1 - i\xi \cdot y\chi_{\hat{B}}(y)]\nu(dy)\Big\}, \tag{2.2}$$

where $\hat{B} = B_1(0)$.

Conversely, any mapping of the form (2.2) is the characteristic function of an infinitely divisible probability measure on $\mathbb{R}^d$.

Because of the equivalence between Lévy process and infinitely indivisible distribution, if X is a Lévy process, we also have the *Lévy-Khintchine formula* for X,

$$E(e^{i\xi\cdot X_t}) = \exp\Big(t\Big\{\int_0^t ib \cdot \xi - \frac{1}{2}\xi \cdot A\xi + \int_{\mathbb{R}^d\backslash\{0\}} [e^{i\xi\cdot y} - 1 - i\xi \cdot y\chi_{\hat{B}}(y)]\nu(dy)\Big\}\Big),$$

for each $t \geqslant 0$, $\xi \in \mathbb{R}^d$, where (b, A, ν) are the *characteristics of* X_1. There are several typical examples of Lévy processes which we can give the explicit characteristics.

Example 2.11. *(The Poisson process, the compensated Poisson process) The Poisson process of intensity $\lambda > 0$ is a Lévy process N taking values in $\mathbb{N} \cup \{0\}$ so that we have*

$$P(N(t) = n) = \frac{(\lambda t)^n}{n!}e^{-\lambda t}$$

for each $n = 0, 1, 2, \cdots$. In this case we have $EN_t = E[N_t]^2 = \lambda t$ for each $t \geqslant 0$, and $\eta(\xi) = \lambda(e^{i\xi} - 1)$.

For later work it is useful to introduce the compensated Poisson process $\tilde{N} = (\tilde{N}_t)_{t\geqslant 0}$ *where each $\tilde{N}_t = N_t - \lambda t$. Note that $E(\tilde{N}_t) = 0$ and $E[\tilde{N}_t]^2 = \lambda t$ for each $t \geqslant 0$.*

Example 2.12. *(Rotationally invariant stable Lévy processes) A* rotationally invariant stable Lévy process *is a Lévy process X where the Lévy symbol is given by*

$$\eta(\xi) = -\sigma^\alpha |\xi|^\alpha,$$

here $\alpha \in (0, 2]$ is the index of stability and $\sigma > 0$. Observe that when $\alpha = 2$, X is the well-known Brownian motion.

Having these basic concepts in mind we are going to simply introduce the pseudo-differential operators which are quite related the Lévy processes introduced above. Actually there is a larger class of processes called Markov processes which usually are introduced and shown to be determined by the associated generator, resolvent and semigroup. Here we focus on Lévy processes only and we introduce two key representations for the generator: first, as a pseudo-differential operator; second, in 'Lévy-Khintchine form', which is the sum of a second-order elliptic differential operator and a (compensated) integral of difference operators.

Let X be a $(\mathcal{F}_t)$-Lévy process in a probability space $(\Omega, \mathcal{F}, P)$. For each $t \geqslant 0$, q_t denote the law of X_t and for each $f \in B_b(\mathbb{R}^d)$, $x \in \mathbb{R}^d$, define

$$(T_t f)(x) =: E[f(X_t + x)] = \int_{\mathbb{R}^d} f(x+y) q_t(dy),$$

then we can get that actually T_t is a Feller semigroup, i.e. T_t is a contracted semigroup in Banach space $C_0(\mathbb{R}^d)$. we have the following important theorem in the analytic study of Lévy processes.

Theorem 2.13. *([2, Theorem 3.3.3]) Let X be a Lévy process with Lévy symbol η and characteristics (b, a, ν), let $(T_t)_{t\geqslant 0}$ be the associated Feller semigroup and A be its infinitesimal generator.*

1. *For each $t \geqslant 0$, $f \in \mathscr{S}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$,*

$$(T_t f)(x) = \frac{1}{(2\pi)^{-d/2}} \int_{\mathbb{R}^d} e^{i\xi\cdot x} e^{t\eta(\xi)} \hat{f}(\xi) d\xi,$$

   so that T_t is a pseudo-differential operator with symbol $e^{t\eta}$.

2. *For each $f \in \mathscr{S}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$,*

$$(Af)(x) = \frac{1}{(2\pi)^{-d/2}} \int_{\mathbb{R}^d} e^{i\xi\cdot x} \eta(\xi) \hat{f} d\xi,$$

   so that A is a pseudo-differential operator with symbol η.

3. *For each $f \in \mathscr{S}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$,*

$$\begin{aligned}(Af)(x) = b \cdot \nabla f(x) &+ \sum_{i,j=1}^{d} a_{ij} \partial_i \partial_j f(x) \\ &+ \int_{\mathbb{R}^d \setminus \{0\}} [f(x+y) - f(x) - y \cdot \nabla f(x) \chi_{\hat{B}}(y)] \nu(dy). \qquad (2.3)\end{aligned}$$

We will now give a number of examples of specific forms of (2.3) corresponding to important examples of Lévy processes.

Example 2.14. *(Standard Brownian motion) Let X be a standard Brrownian motion in $\mathbb{R}^d$. Then X has characteristics $(0, I, 0)$, and so we see from (2.3) that*

$$A = \frac{1}{2} \sum_{i=1}^{d} \partial_i^2 = \frac{1}{2} \Delta,$$

where Δ is the usual Laplacian operator.

Example 2.15. *(Brownian motion with drift) Let X be a Brownian motion with drift in $\mathbb{R}^d$. Then X has characteristics $(b, a, 0)$ and A is a diffusion operator of the form*

$$A = b \cdot \nabla + \sum_{i,j=1}^{d} a_{ij} \partial_i \partial_j .$$

Example 2.16. *(Rotationally invariant stable processes) Let X be a rotationally invariant stable process of index $\alpha \in (0, 2)$. Its symbol is given by $\eta(\xi) = -|\xi|^\alpha$ for all $u \in \mathbb{R}^d$, then*

$$A = -(-\Delta)^{\frac{\alpha}{2}},$$

i.e. the fractional Laplacian operator.

2.3 Strong solutions, weak solutions and martingale solutions to SDEs

In order to make the definition of solutions to the SDEs clear, we recall some classical terminology. Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with a filtration $\mathcal{F}_t$ that satisfies the usual conditions. Let W be an d-dimensional standard Brownian motion and N an independent Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d \backslash \{0\}$ with associated compensator $\tilde{N}$ and intensity measure ν, where we assume that ν is a Lévy measure. We always assume that W and N are independent of $\mathcal{F}_0$. We consider the following SDE: for $t \geqslant 0$,

$$X_t = x + \int_0^t b(s, X_{s-}) ds + \int_0^t \sigma(s, X_{s-}) dW_s + \int_0^t \int_{|z|<1} g(s, X_{s-}, z) \tilde{N}(ds, dz) + \int_0^t \int_{|z|\geqslant 1} g(s, X_{s-}, z) N(ds, dz), \quad t \geqslant 0, \tag{2.4}$$

Here the mappings $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$, $g : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are all assumed to be measurable. Then

1. *weak existence* holds for SDE (2.4) if one can construct a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$, and an adapted Brownian motion W and an adapted Poisson measure N and an adapted process X on this space which satisfies SDE (2.4).

2. *Uniqueness in law* holds if every solution X to (2.4), possibly on different probability space, has the same law.

3. *Strong existence* means that one can find a solution to (2.4) on any given filtered probability space equipped with any given adapted Brownian motion.

4. *pathwise uniqueness* means that, on any given filtered probability space equipped with any given Brownian motion and Poisson measure N, any two solutions to (2.4) with the same given $\mathcal{F}_0$-measurable initial condition x coincide.

Notice that, in contrast to strong solutions, where the noise is prescribed in advance, for weak solutions the construction of the noise is part of the problem. Finding weak solution to SDE is intimately related to martingale problems. For linear operator $\mathcal{L}_t$ defined as

$$\begin{aligned}\mathcal{L}_t f(x) := b(t,x)\cdot\nabla f(x) + \frac{1}{2}\sum_{i,j=1,k=1}^{d} \sigma_{ik}(t,x)\sigma_{jk}(t,x)\partial_i\partial_j f(x) \\ + \int_{|z|<1}\Big(f(x+g(t,x,z)) - f(x) - g(t,x,z)\cdot\nabla f(x)\Big)\nu(dz) \\ + \int_{|z|\geqslant 1}\Big(f(x+g(t,x,z) - f(x))\Big)\nu(dz). \end{aligned} \tag{2.5}$$

The definition about the martingale solution goes as following.

Definition 2.17. *A probability measure P on $(\mathcal{C}[0,\infty), \mathcal{B}(\mathcal{C}[0,\infty)^d))$ which is cádlág, under which*

$$M_t^f = f(w(t)) - f(w(0)) - \int_0^t (\mathcal{L}_s f)(w)ds, \quad 0 \leqslant t < \infty, \tag{2.6}$$

is a continuous, local martingale for every $f \in \mathcal{C}^2(\mathbb{R}^d)$, is called a martingale solution to the local martingale problem *associated with $\mathcal{L}_t$.*

2.4 Estimates of the fundamental solutions to second order parabolic equations

Because of the crucial role of the estimate of fundamental solution of parabolic equation in Chapter 5, in this subsection we collect the results that we will use. For the detailed discussion we refer to [40, IV] and [55].

First we consider the Cauchy problem with terminal data for equation in the domain $[0,T]\times\mathbb{R}^d$.

Let there be given in the domain $Q_T := (0,T)\times\mathbb{R}^d$ a parabolic operator $\mathcal{L}(x,t,\frac{\partial}{\partial x},\frac{\partial}{\partial t})$ defined as

$$\mathcal{L}(x,t,\frac{\partial}{\partial x},\frac{\partial}{\partial t})u = \frac{\partial u}{\partial t} - \sum_{i,j=1}^{d} a_{ij}u(t,x)\frac{\partial^2 u}{\partial x_i \partial x_j} + b(x,t)\cdot\nabla u + c(t,x)u$$

with coefficients $(a_{ij})_{1\leqslant i,j\leqslant d}$, b and c belonging to the Hölder space $\mathcal{C}^{1/2,1}(Q_T)$, where the Hölder space $\mathcal{C}^{1/2,1}(Q_T)$ is the Bananch space of functions $u(t,x)$ that are continuous in Q_T with a finite form

$$|u|_{\mathcal{C}^{1/2,1}(Q_T)} = \langle u\rangle_{Q_T} + \max_{(t,x)\in Q_T}|u(t,x)|$$

where

$$\langle u\rangle_{Q_T} = \langle u\rangle_{x,Q_T} + \langle u\rangle_{t,Q_T},$$

$$\langle u\rangle_{x,Q_T} = \sup_{\substack{(x,t),(x',t)\in\overline{Q_T}\\ x\neq x'}} \frac{|u(t,x)-u(t,x')|}{|x-x'|},$$

$$\langle u\rangle_{t,Q_T} = \sup_{\substack{(x,t),(x',t)\in\overline{Q_T}\\ t\neq t'}} \frac{|u(t,x)-u(t,x')|}{|t-t'|^{1/2}}.$$

We assume that this operator is uniformly elliptic, i.e. there exist two positive numbers δ_1 and δ_2 such that for all $(t,x)\in Q_T$

$$\delta_1\xi^2 \leqslant \sum_{i,j=1}^{d} a_{ij}(t,x)\xi_i\xi_j \leqslant \delta_2\xi^2, \quad \forall\xi=(\xi_1,\cdots,\xi_d)\in\mathbb{R}^d.$$

We consider the following Cauchy problem on Q_T

$$\begin{cases} \mathcal{L}(x,t,\frac{\partial}{\partial x},\frac{\partial}{\partial t})u(t,x) = 0, \\ \qquad\qquad\qquad u(0,x) = \phi(x). \end{cases} \tag{2.7}$$

[40, Theorem 5.1] implies that (2.7) has a unique solution $u\in\mathcal{C}^{1,2}(Q_T)$ (actually u can be smoother but more lengthy and such regularity is enough for use), i.e. $u(t,x)$ is $1-$order differentiable with respect to $t\in[0,T]$ with $|\frac{\partial u}{\partial t}|$ bounded on Q_T and is $2-$orders differentiable with respect to $x\in\mathbb{R}^d$ with $|\frac{\partial u}{\partial x_i}|, 1\leqslant i\leqslant d$, and $|\frac{\partial^2 u}{\partial x_i\partial x_j}|, 1\leqslant i\leqslant j\leqslant d$, bounded on Q_T, and these bounds are controlled by $\sup_{x\in\mathbb{R}^d}|\phi(x)|$.

We call function $Z(t,x;s,y): Q_T\times Q_T\to\mathbb{R}$ a fundamental solution if Z satisfies the equation

$$\mathcal{L}(x,t,\frac{\partial}{\partial x},\frac{\partial}{\partial t})Z(t,x;s,y) = \delta(x-y)\delta(t-s)$$

and is bounded for $|x|\to\infty$. The function Z plays the same important role for the operator $\mathcal{L}$ as the function $g(t,x;s,y)=\frac{1}{(4\pi|t-s|)^{d/2}}\exp(\frac{|x-y|^2}{4|t-s|})$ for the heat operator $\mathcal{H}=\frac{\partial}{\partial t}-\Delta$.

If ϕ is continuous, [40, 14.1] says the solution to (2.7) can be written in the form of a potential with kernel Z:

$$u(t,x) = \int_{\mathbb{R}^d} Z(t,x;0,y)\phi(y)dy. \tag{2.8}$$

Besides, the following estimates of Z were obtained in [40, 13.1,13.2], which says: for $2m+n\leqslant 2$, $t>s$

$$|D_t^m D_x^n Z(t,x;s,y)| \leqslant C(t-s)^{-\frac{d+2m+n}{2}}\exp\Big(-C\frac{|x-y|^2}{t-s}\Big), \tag{2.9}$$

for $2m+n=2(i.e. m=0, n=2$ and $m=1, n=0)$, $0\leqslant\gamma\leqslant 1$, $0\leqslant\beta\leqslant 1$, $t>s$,

$$|D_t^m D_x^n Z(t,x;s,y)-D_t^m D_x^n Z(t,x';s,y)| \leqslant C\Big[|x-x'|^{\gamma}(t-s)^{-\frac{d+2+\gamma}{2}}+|x-x'|^{\beta}(t-s)^{-\frac{d+2-1+\beta}{2}}\Big]\exp\Big(-\frac{|x-y|^2}{t-s}\Big), \quad (2.10)$$

and for $2m+n=1,2$ and $t>t'>s$,

$$|D_t^m D_x^n Z(t,x;s,y)-D_t^m D_x^n Z(t',x;s,y)| \leqslant C\Big[(t-t')(t-s)^{-\frac{d+2m+s+2}{2}}+(t-t')^{\frac{2-2r-s+\alpha}{2}}(t'-s)^{-\frac{d+2}{2}}\Big]\exp\Big(-\frac{|x-y|^2}{t-s}\Big). \quad (2.11)$$

For fixed time $T\in[0,\infty)$, if we denote $v(T-t,x)=u(t,x)$, then v solves the following backward equation

$$\begin{cases} \mathcal{L}'(x,t,\frac{\partial}{\partial x},\frac{\partial}{\partial t})v(t,x)=0, \\ \qquad\qquad v(T,x)=\phi(x). \end{cases} \quad (2.12)$$

Where $\mathcal{L}'$ is defined as

$$\mathcal{L}'(x,t,\frac{\partial}{\partial x},\frac{\partial}{\partial t})v=\frac{\partial v}{\partial t}+\sum_{i,j=1}^{d}a_{ij}(t,x)\frac{\partial^2 v}{\partial x_i\partial x_j}+b(x,t)\cdot\nabla v+c(t,x)v.$$

Corresponding to (2.8), the solution v to the backward equation (2.12) can be represented as

$$v(t,x)=\int_{\mathbb{R}^d}Z(T,y;t,x)\phi(y)dy.$$

In this case the estimates (2.9), (2.10) and (2.11) still holds for the equation (2.12) with the similar form, i.e. for $2m+n\leqslant 2$, $t<T$

$$|D_t^m D_x^n Z(T,y;t,x)|\leqslant C(T-t)^{-\frac{d+2m+n}{2}}\exp\Big(-C\frac{|x-y|^2}{T-t}\Big),$$

for $2m+n=2$ (i.e. $m=0, n=2$ and $m=1, n=0$), $0\leqslant\gamma\leqslant 1$, $0\leqslant\beta\leqslant 1$, $t<T$,

$$|D_t^m D_x^n Z(T,y;t,x)-D_t^m D_x^n Z(T,y;t,x')| \leqslant C\Big[|x-x'|^{\gamma}(T-t)^{-\frac{d+2+\gamma}{2}}+|x-x'|^{\beta}(T-t)^{-\frac{d+2-1+\beta}{2}}\Big]\exp\Big(-\frac{|x-y|^2}{T-t}\Big),$$

and for $2m+n=1,2$ and $T>t>t'$,

$$|D_t^m D_x^n Z(T,y;t,x)-D_t^m D_x^n Z(T,y;t',x)| \leqslant C\Big[(t-t')(T-t)^{-\frac{d+2m+s+2}{2}}+(t-t')^{\frac{2-2r-s+\alpha}{2}}(T-t')^{-\frac{d+2}{2}}\Big]\exp\Big(-\frac{|x-y|^2}{T-t}\Big).$$

As the second part of this subsection, we consider the first boundary problem to the following parabolic equation on the cylindrical domain $\overline{Q^{r^2,r}}$ with surface $\partial Q^{r^2,r} := ((0,r^2)\times\partial B_r)\cup(\{r^2\}\times B_r)$ for $r\in(0,1]$ assuming that f is a continuous function on $\partial Q^{r^2,r}$:

$$\begin{cases} \mathcal{L}u(t,x) = D_t u(t,x) + \frac{1}{2}\sum_{i,j=1}^{d}\partial_i(a_{ij}(t,x)\partial_j u(t,x)) = 0 \qquad \textit{on} \quad Q^{r^2,r}, \\ u(t,x) = f(t,x) \qquad \textit{on} \quad \partial Q^{r^2,r}, \end{cases} \tag{2.13}$$

where $(a_{ij})_{1\leqslant i,j\leqslant d}$ is assumed to be real, symmetric and uniformly elliptic, i.e. for some $\mu\geqslant 1$, for all $(t,x)\in Q^{r^2,r}$ and all $\xi\in\mathbb{R}^d$, $1/\mu\xi^2\leqslant\sum_{i,j=1}^d a_{ij}(t,x)\xi_i\xi_j\leqslant\mu\xi^2$, with μ-Lipschitz coefficients with respect to the parabolic distance, i.e. for all $1\leqslant i,j\leqslant d$

$$|a_{ij}(t,x)-a_{ij}(s,y)|\leqslant\mu(|x-y|\vee|t-s|^{1/2}).$$

[55, Corollary 3.2] says there exists a Possion kernel $p(t,x;s,y) : Q^{r^2,r}\times\partial Q^{r^2,r}\to\mathbb{R}_+$ such that the potential

$$u(t,x) = \int_{\partial Q^{r^2,r}} p(t,x;s,y)f(s,y)dS(s,y)$$

represents the solution to (2.13), where dS denotes the surface measure on $\partial Q^{r^2,r}$. And [55, Theorem 3.1] says that this Possion kernel p satisfies the following estimates: there exists a constant $k>0$ depending only on d, μ, $Q^{r^2,r}$ such that

$$\frac{1}{k}|t-s|^{-\frac{d+1}{2}}\exp(C\frac{|x-y|^2}{|t-s|})\leqslant p(t,x;s,y)\leqslant k|t-s|^{-\frac{d+1}{2}}\exp(C\frac{|x-y|^2}{|t-s|}),$$

for all $(t,x)\in Q^{r^2,r}$ and $(s,y)\in\partial Q^{r^2,r}$.

2.5 Main methods

We are going to introduce the main methods applied in this thesis.

2.5.1 Zvonkin's transformation

Originally invented in the paper [79], *Zvonkin's transformation* is one of the main tool to prove existence and uniqueness of the solution in most of the papers mentioned above and it also plays a crucial role in our work. We now give the introduction about the idea behind. Let $\mathscr{L}_2^\sigma$ be the second order differential operator related to diffusion coefficients σ which is defined as

$$\mathscr{L}_2^\sigma f(t,x) := \frac{1}{2}\sum_{i,j=1,k=1}^{d}\sigma_{ik}(t,x)\sigma_{jk}(t,x)\partial_i\partial_j f(t,x), \quad f\in\mathcal{C}_c^\infty(\mathbb{R}^{d+1}),$$

let $\mathscr{L}_1^b$ be the first order differential operator related to drift coefficients b as

$$\mathscr{L}_1^b f(t,x) = b(t,x)\cdot\nabla f(t,x), \quad f\in\mathcal{C}_c^\infty(\mathbb{R}^{d+1}),$$

and let $\mathscr{L}_\nu^g$ be the nonlocal operator associated with the jump coefficient g whereas

$$\mathscr{L}_\nu^g f(t,x) := \int_{|z|<1}\Big(f(t,x+g(t,x,z)) - f(t,x) - g(t,x,z)\cdot\nabla f(t,x)\Big)\nu(dz) + \int_{|z|\geqslant 1}\Big(f(t,x+g(t,x,z) - f(t,x))\Big)\nu(dz). \tag{2.14}$$

Consider the following equation

$$\partial_t u + (\mathscr{L}_2^\sigma + \mathscr{L}_1^b + \mathscr{L}_\nu^g)u = 0, \quad u(T,x) = 0\in\mathbb{R}^d, \tag{2.15}$$

if this equation has a regular enough solution u such that for each $t\in[0,T]$, the map $\Phi(x) := x + u(x)$ forms a $\mathcal{C}^2-$diffeomorphism on $\mathbb{R}^d$, then by applying Itô's formula to the solution $(X_t)_{t\geqslant 0}$ to (1.1)

$$\begin{aligned}\Phi(X_t) = \Phi(X_0) &+ \int_0^t \nabla\Phi(X_s)\sigma(s,X_s)dW_s\\ &+ \int_0^t\int_{|z|<1}(\Phi(X_{s-}+g(s,X_{s-},z)) - \Phi(X_{s-}))\tilde{N}(ds,dz)\\ &+ \int_0^t\int_{|z|\geqslant 1}(\Phi(X_{s-}+g(s,X_{s-},z)) - \Phi(X_{s-}))N(ds,dz).\end{aligned}$$

If we denote $Y_t := \Phi(X_t), t\geqslant 0$ and

$$\hat{\sigma}(t,y) := (\nabla\Phi\cdot\sigma(t,\cdot))\circ\Phi^{-1}(y), \quad \hat{g}(t,y,z) := \Phi(\Phi^{-1}(y) + g(t,\Phi^{-1}(y)z)) - y,$$

then Y_t satisfies the following new SDE *without irregular* drift:

$$\begin{aligned}Y_t = \Phi(x) + \int_0^t\hat{\sigma}(s,Y_s)dW_s &+ \int_0^t\int_{|z|<1}\hat{g}(s,Y_{s-},z)\tilde{N}(ds,dz)\\ &+ \int_0^t\int_{|z|\geqslant 1}\hat{g}(s,Y_{s-},z)N(ds,dz), \quad t\geqslant 0.\end{aligned} \tag{2.16}$$

Hence it is equivalent to solve SDE (1.1) via solving SDE (2.16) instead, which has no irregular drift term and the coefficients of the noise term could be a bit continuous because of the second order regularization effect of equation (2.15). Then the main task is to solve equation (2.15) such that Φ has the desired properties. To this purpose a key step is to show the following *Krylov's estimate*: for any solution Y, and any $T>0$ and $f\in L^q_{loc}([0,\infty);L^p(\mathbb{R}^d))$,

$$E\Big(\int_0^T|f(t,Y_t)|dt\Big) \leqslant c\Big(\int_0^T\Big(\int_{\mathbb{R}^d}|f(t,x)|^p dx\Big)^{q/p}dt\Big)^{1/q}. \tag{2.17}$$

When $g=0$, such estimate was established in [37] by Krylov and Röckner when $\sigma=\mathbb{I}_d$ and $\frac{d}{p}+\frac{2}{q}<2$. Zhang and other authors showed it for the case when σ is nonconstant matrix-valued functions in [73, 70, 76, 77] and references therein. In [69, 58, 5] they obtained the estimate for the SDE with jump noise.

2.5.2 Density of the solution to SDE driven by Lévy noise

The idea is from [15] to prove the existence and regularity of density to jump type SDEs in our work. The method therein is to apply the following crucial lemma:
Define, for $f:\mathbb{R}^d\to\mathbb{R}$, for x, $h\in\mathbb{R}^d$ and $n\geqslant 1$,

$$(\Delta_h^1 f)(x)=f(x+h)-f(x),\quad (\Delta_h^n f)(x)=\Delta_h^1(\Delta_h^{n-1}f)(x),$$

$\mathcal{M}(\mathbb{R}^d)$ denotes the set of probability measures on $\mathbb{R}^d$. The lemma says

Lemma 2.18. *([15, Lemma 2.1]) Let $\rho\in\mathcal{M}(\mathbb{R}^d)$. Assume that there are $0<\eta<a<1$, $n\geqslant 1$ and a constant K such that for all $\phi\in\mathcal{C}^\eta(\mathbb{R}^d)$, all $h\in\mathbb{R}^d$ with $|h|\leqslant 1$,*

$$\|\int_{\mathbb{R}^d}\Delta_h^n\phi(x)\rho(dx)\|\leqslant K\|\phi\|_{\mathcal{C}^\eta}|h|^a.$$

Then ρ has a density in $B^{a-\eta}_{1,\infty}(\mathbb{R}^d)$ (Besov space) and $\|\rho\|_{B^{a-\eta}_{1,\infty}}\leqslant\rho(\mathbb{R}^d)+C_{d,a,\eta,n}K$.

Then the strategy to apply this lemma to jump type SDE

$$X_t=x+\int_0^t g(X_{s-})dL_s+\int_0^t b(X_s)ds,\quad t\geqslant 0,$$

is the following:

- For $\epsilon\in(0,t)$, consider

$$X_t^\epsilon=X_{t-\epsilon}+\epsilon b(X_{t-\epsilon})+g(X_{t-\epsilon})(L_t-L_{t-\epsilon}).$$

- Study the error $E[|X_t-X_t^\epsilon|^\eta]$ for $\eta>0$. We get something like $E[|X_t-X_t^\epsilon|^\eta]\leqslant C\epsilon^\gamma$ with γ depending on η,α, and on the Hölder regularity of the coefficients g and b.

- Conditionally on $X_{t-\epsilon}$, X_t^ϵ has an infinitely divisible distribution, for which many known results are available. We can get the bound of any derivatives of the density $f_{X_t^\epsilon}$ in $L^1(\mathbb{R}^d)$, which will explodes when $\epsilon\to 0$ but the rate of the growth is controlled precisely: we obtain that $\|D^n f_{X_t^\epsilon}\|_{L^1(\mathbb{R}^d)}\leqslant\epsilon^{-n/\alpha}$.

- Use the discrete integration by part:

$$E(\Delta_h^n\phi(X_t^\epsilon))=\int_{\mathbb{R}^d}\Delta_h^n\phi(x)f_{X_t^\epsilon}(x)dx=\int_{\mathbb{R}^d}\phi(x)\Delta_{-h}^n f_{X_t^\epsilon}(x)dx.$$

  To obtain

$$|E(\Delta_h^n\phi(X_t^\epsilon))|\leqslant\|\phi\|_{L^\infty}\|D^n f_{X_t^\epsilon}\|_{L^1(\mathbb{R}^d)}|h|^n\leqslant\|\phi\|_{L^\infty}\epsilon^{-n/\alpha}|h|^n.$$

- Last step is to write

$$\begin{aligned}|E(\Delta_h^n \phi(X_t))| &\leqslant |E(\Delta_h^n \phi(X_t^\epsilon))| + |E(\Delta_h^n \phi(X_t)) - E(\Delta_h^n \phi(X_t^\epsilon))| \\ &\leqslant C\|\phi\|_{L^\infty} \epsilon^{-n/\alpha} |h|^n + C\|\phi\|_{\mathcal{C}^\eta} E|X_t - X_t^\epsilon|^\eta.\end{aligned}$$

  For each h, choose ϵ suitable enough to end the results like

$$\| \int_{\mathbb{R}^d} \Delta_h^n \phi(x) f_{X_t}(x) dx \| = |E(\Delta_h^n \phi(X_t))| \leqslant \|\phi\|_{\mathcal{C}^\eta} |h|^\delta,$$

  for some δ depending on α, on the Hölder regularity of the coefficients g and b, and n, η, about which η is suitable enough to guarantee that the above lemma could be applied.

3 Existence and Uniqueness of a global strong solution to an SDE driven by continuous noise in mixed-norm Lebesgue spaces on $Q = [0, \infty) \times \mathbb{R}^d$

3.1 Preliminaries and main results

Consider the following SDE in $\mathbb{R}^d$:

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t, \quad X_0 = x, \tag{3.1}$$

where $d \geqslant 1$, $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ is a Borel measurable function, and $(W_t)_{t\geqslant 0}$ is a standard Brownian motion defined on some probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbb{P})$. The remarkable result from N. V. Krylov and M. Röckner [37] shows that if

$$b \in L^q_{loc}(\mathbb{R}_+; L^p_{loc}(\mathbb{R}^d)) \quad \text{with} \quad p, q \in (2, \infty) \text{ and } d/p + 2/q < 1, \tag{3.2}$$

then for each $x \in \mathbb{R}^d$, there exists a unique strong solution $(X_t)_{t\geqslant 0}$ for SDE (3.1) up to the explosion time. Later, X. Zhang [73] extend this result to SDEs driven by multiplicative noise

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t, \quad X_0 = x \tag{3.3}$$

under the assumptions that σ is a bounded, uniformly elliptic matrix-valued function which is uniformly continuous in x locally uniformly with respect to t, and

$$|b|, |\nabla\sigma| \in L^q_{loc}(\mathbb{R}_+; L^p_{loc}(\mathbb{R}^d))$$

with $p, q \in (1, \infty)$ satisfying (3.2). Here and below, ∇ denotes the weak derivative with respect to x variable. Note that when $\sigma \equiv 0$ in (3.3), the corresponding deterministic ordinary differential equation is far from being well-posed under the above condition on the drift coefficient. This is known as the regularization effect of noises, we refer to [28] for a comprehensive overview. From then on, there are increasing interests of studying the strong well-posedness as well as properties of the unique strong solution for SDE (3.3) with singular coefficients, see e.g. [21, 49, 65, 69, 75] and references therein.

However, there seems to be one non-uniform place in the above mentioned results concerning the strong well-posedness for SDE (3.1) and (3.3): the conditions imposed on the coefficients will not be consistent. Let us specify this by the following example: consider the following SDE in $\mathbb{R}^2$:

$$\begin{cases} \mathrm{d}X^1_t = b_1(t, X^1_t)\mathrm{d}t + \mathrm{d}W^1_t, & X^1_0 = x_1 \in \mathbb{R}, \\ \mathrm{d}X^2_t = b_2(t, X^2_t)\mathrm{d}t + \mathrm{d}W^2_t, & X^2_0 = x_2 \in \mathbb{R}. \end{cases} \tag{3.4}$$

If we denote $x := (x_1, x_2)^* \in \mathbb{R}^2$, $X_t := (X^1_t, X^2_t)^*$, $W_t := (W^1_t, W^2_t)^*$, and define the vector field

$$b(t, x) := \begin{pmatrix} b_1(t, x_1) \\ b_2(t, x_2) \end{pmatrix}.$$

Then SDE (3.4) can be rewritten as

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t, \quad X_0 = x \in \mathbb{R}^2. \tag{3.5}$$

According to the above mentioned results, we need to assume

$$b \in L^q_{loc}(\mathbb{R}_+; L^p_{loc}(\mathbb{R}^2)) \quad \text{with} \quad 2/p + 2/q < 1$$

to ensure the well-posedness of SDE (3.5). This in particular means that we need

$$b_1, b_2 \in L^q_{loc}(\mathbb{R}_+; L^p_{loc}(\mathbb{R}^1)) \quad \text{with} \quad 2/p + 2/q < 1. \tag{3.6}$$

On the other hand, the two-dimensional SDE (3.4) can also be viewed as single equations for X^1_t and X^2_t their-self, because X^1_t and X^2_t are not involved together in the equation. From this point of view, SDE (3.4) can be well-posed under the condition that

$$b_1, b_2 \in L^q_{loc}(\mathbb{R}_+; L^p_{loc}(\mathbb{R}^1)) \quad \text{with} \quad 1/p + 2/q < 1, \tag{3.7}$$

which do not coincides with (3.6). We point out that such ununify will always appear when we consider SDEs in multi-dimensional, and especially for degenerate noise cases and multi-scales models involving at least slow and fast phase variables, see e.g. [22, 68].

The results of this chapter is based on the joint work [44] from author and X. Long. The main aim of this work is to get rid of the above unreasonableness by studying SDE (3.3) with coefficients in general mixed-norm spaces (cf. [44]). To this end, let $\mathbf{p} = (p_1, \cdots, p_d) \in [1,\infty)^d$ be a multi-index, we denote by $L^{\mathbf{p}}(\mathbb{R}^d)$ the space of all measurable functions on $\mathbb{R}^d$ with norm

$$\|f\|_{L^{\mathbf{p}}} := \left(\int_{\mathbb{R}} \cdots \left(\int_{\mathbb{R}} |f(x_1, ..., x_d)|^{p_1} \mathrm{d}x_1 \right)^{\frac{p_2}{p_1}} \mathrm{d}x_2 \right)^{\frac{p_3}{p_2}} \cdots \mathrm{d}x_d \Bigg)^{\frac{1}{p_d}} < \infty.$$

When $p_i = \infty$ for some $i = 1, \cdots, d$, the norm is taken as supreme with respect to the corresponding variable. Notice that the order is important when we take above integrals. If we permute the p_is, then increasing the order of p_i gives the smallest norm, while by decreasing the order gives the largest norm.

Our main result in this chapter is as follows.

Theorem 3.1. *Assume that for some $p_1, \cdots, p_d, q \in (2,\infty)$ and every $T > 0$,*

$$|b|, |\nabla\sigma| \in L^q([0,T]; L^{\mathbf{p}}(\mathbb{R}^d)) \quad \text{with} \quad \frac{2}{q} + \frac{1}{p_1} + \cdots + \frac{1}{p_d} < 1, \tag{3.8}$$

and σ is uniformly continuous in $x \in \mathbb{R}^d$ uniformly with respect to $t \in [0,T]$, and there exist positive constants δ_1 and δ_2 such that for all $(t,x) \in [0,T] \times \mathbb{R}^d$,

$$\delta_1|\xi|^2 \leqslant |\sigma^*(t,x)\xi|^2 \leqslant \delta_2|\xi|^2, \quad \forall \xi \in \mathbb{R}^d.$$

Then for any $(\mathcal{F}_t)-$stopping time τ and $x \in \mathbb{R}^d$, there exists a unique $(\mathcal{F}_t)-$adapted continuous solution $(X_t)_{t\geqslant 0}$ such that

$$P\left\{\omega : \int_0^T |b(r, X_r(\omega))|dr + \int_0^T |\sigma(r, X_r(\omega))|^2 dr < \infty, \forall T \in [0, \tau(\omega))\right\} = 1, \tag{3.9}$$

and

$$X_t = x + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r, \quad \forall t \in [0, \tau), a.s. \tag{3.10}$$

which means that if there is another $(\mathcal{F}_t)-$adapted continuous stochastic process $(Y_t)_{t\geqslant 0}$ also satisfying (3.9) and (3.10), then

$$P\{\omega : X_t(\omega) = Y_t(\omega), \forall t \in [0, \tau)\} = 1.$$

Moreover, for almost all ω and all $t \geqslant 0$, $x \to X_t(\omega, x)$ is a homeomorphism on $\mathbb{R}^d$ and for any $t > 0$ and bounded measurable function ψ, x, $y \in \mathbb{R}^d$,

$$|E\psi(X_t(x)) - E\psi(X_t(y))| \leqslant C_t \|\psi\|_\infty |x - y|, \tag{3.11}$$

where $C_t > 0$ satisfies $\lim_{t\to 0} C_t = +\infty$.

Remark 3.2. *i) The advantage of (3.8) lies in the flexible for integrability of the coefficients in different directions. More precisely, it allows the integrability of the coefficients to be small in some directions by taking the integrability index large for the other directions. With this condition, the problem of the tricky example mentioned before will not appear since we can take another index to be ∞. To be more specific, according to Theorem 3.1 there exits a unique $(\mathcal{F}_t)-$adapted solution to the SDE (3.4) if $b = (b_1, b_2) \in L^q([0, T], L^{\mathbf{p}}(\mathbb{R}^2))$ with $2/q + 1/p_1 + 1/p_2 < 1$. That is to say, $b_1, b_2 \in L^q([0, T], L^{\mathbf{p}}(\mathbb{R}^2))$ with $2/q + 1/p_1 + 1/p_2 < 1$, by the definition of the mixed-norm Lebesgue space, for $b_1 \in L^q([0, T], L^{\mathbf{p}}(\mathbb{R}^2))$ we can take $(p_1, p_2) = (p_1, \infty)$ since b_1 is only defined on $\mathbb{R}$, the same with b_2. Hence the integrability conditions of b_1 and b_2 are $2/q + 1/p_1 < 1$ and $2/q + 1/p_2 < 1$ respectively, which coincides with (3.7).*

ii) As mentioned in [35], the necessity of mixed-norm spaces arises when the physical processes have different behavior in each component. In view of (3.8), it reflects the classical fact that the integrability of time variable and space variable have the ratio 1:2. Meanwhile, the integrability of space variables in each direction is the same, which is natural because the noise is non-degenerate. Such mixed-norm spaces will be more important when studying SDEs with degenerate noise. This will be our future works.

Remark 3.3. *In above theorem the condition $p_1, \cdots, p_d, q \in (2, \infty)$ is automatically fulfilled when $d \geqslant 2$ since we also assume $1/p + \cdots + 1/p_d + 2/q < 1$. When $d = 1$, we can refer to the result from Engelbert and Schmidt [18] to obtain the existence and uniqueness of a strong solution to homogeneous SDE on $\mathbb{R}^d$. It proved that if $\sigma(x) \neq 0$ for all $x \in \mathbb{R}$ and $b/\sigma^2 \in L^1_{loc}(\mathbb{R})$, and there exists a constant $C > 0$ such that*

$$|\sigma(x) - \sigma(y)| \leqslant C\sqrt{|x - y|}, \quad x, y \in \mathbb{R},$$

$$|b(x)| + |\sigma(x)| \leqslant C(1 + |x|),$$

then there exists a unique $(\mathcal{F}_t)-$ *adapted process* $(X_t)_{t\geqslant 0}$ *such that the following SDE hold:* $dX_t = b(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x \in \mathbb{R}^d$.

Now, let us specify the proof briefly. The key tool to prove our main result is the $\mathbb{L}^q_{\mathbf{p}}$-maximal regularity estimate for the following second order parabolic PDEs on $[0, T] \times \mathbb{R}^d$:

$$\partial_t u(t, x) = \mathscr{L}^a_2 u(t, x) + \mathscr{L}^b_1 u(t, x) + f(t, x), \quad u(T, x) = 0, \tag{3.12}$$

where $\mathscr{L}^a_2 + \mathscr{L}^b_1$ is the infinitesimal operator of process $(X_t)_{t\geqslant 0}$, i.e.,

$$\mathscr{L}^a_2 u(t, x) := \frac{1}{2} a^{ij}(t, x)\partial_{ij} u(t, x), \quad \mathscr{L}^b_1 u(t, x) := b^i(t, x)\partial_i u(t, x)$$

with $a(t, x) = (a^{ij}(t, x)) := \sigma\sigma^*(t, x)$, and ∂_i denotes the i-th partial derivative respect to x. Here we use Einstein's convention that the repeated indices in a product will be summed automatically. More precisely, for any $q \in (1, \infty)$ and $\mathbf{p} \in (1, \infty)^d$, we want to establish the following estimate:

$$\|\partial^2 u\|_{\mathbb{L}^q_{\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \tag{3.13}$$

see Section 3.2 for the precise definition of $\mathbb{L}^q_{\mathbf{p}}(T)$. Notice that when $p_1 = \cdots = p_d = q$, it is a standard procedure to prove (3.13) by the classical freezing coefficient argument (cf. [75]). While for general $q \in (1, \infty)$ and $\mathbf{p} \in (1, \infty)^d$, it seems to be non-trivial. When a^{ij} is independent of x and $p_1 = \cdots = p_d$, (3.13) was first proved by Krylov in [35]. In the spatial dependent case, Kim [38] showed (3.13) only for $p_1 = \cdots = p_d \leqslant q$. This was recently generalized to general $p_1 = \cdots = p_d > 1$ and $q > 1$ in [67] by a duality method. We shall further develop the dual argument used in [67], and combing with the interpolation technique, to prove (3.13) for mixed-norms even in the space variable. The main result is provide by Theorem 3.4, which should be of independent interest in the theory of PDEs.

This chapter is organized as follows: In section 3.2, we study the maximal regularity estimate for second order parabolic equations. In section 3.3, we prove the Krylov's estimate. The existence and uniqueness of the global strong solution to SDE are shown in section 3.4.

3.2 Regularity estimates for parabolic type partial differential equations

Fix $T > 0$ and let $\mathbb{R}^{d+1}_T := [0, T] \times \mathbb{R}^d$. This section is devoted to study the parabolic equation (3.12) on $\mathbb{R}^{d+1}_T$ in general mixed-norm spaces. Let us first introduce some spaces and notations.

For $\mathbf{p} = (p_1, ..., p_d) \in [1, \infty)^d$, let $W^2_{\mathbf{p}}(\mathbb{R}^d)$ be the second order Sobolev space which consists of functions $f \in L^{\mathbf{p}}(\mathbb{R}^d)$ such that the second order weak derivative $\nabla^2 f \in$

$L^{\mathbf{p}}(\mathbb{R}^d)$. For $q \in [1,\infty)$ and any $S < T$, denote by $\mathbb{L}^q_{\mathbf{p}}(S,T) := L^q([S,T]; L^{\mathbf{p}}(\mathbb{R}^d))$. For simplicity, we will write $\mathbb{L}^q_{\mathbf{p}}(T) := \mathbb{L}^q_{\mathbf{p}}(0,T)$, and $\mathbb{L}^\infty(T)$ consists of functions satisfying

$$\|f\|_{\mathbb{L}^\infty(T)} := \sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} |f(t,x)| < +\infty.$$

We also introduce

$$\mathbb{W}^q_{2,\mathbf{p}}(T) := L^q\big([0,T]; W^2_{\mathbf{p}}(\mathbb{R}^d)\big),$$

and the space $\mathscr{W}^q_{2,\mathbf{p}}(T)$ consisting of functions $u = u(t)$ on $[0,T]$ with values in the space of distributions on $\mathbb{R}^d$ such that $u \in \mathbb{W}^q_{2,\mathbf{p}}(T)$ and $\partial_t u \in \mathbb{L}^q_{\mathbf{p}}(T)$. Besides, for $\gamma \geqslant 0$, let $H^\gamma_{\mathbf{p}} := (1-\Delta)^{-\gamma/2}(L^{\mathbf{p}}(\mathbb{R}^d))$ be the usual Bessel potential space with norm

$$\|f\|_{H^\gamma_{\mathbf{p}}} := \|(1-\Delta)^{\gamma/2} f\|_{L^{\mathbf{p}}(\mathbb{R}^d)},$$

and $(1-\Delta)^{-\gamma/2} f$ is defined through Fourier transform

$$(1-\Delta)^{\gamma/2} f := \mathcal{F}^{-1}((1+|\cdot|^2)^{\gamma/2}\mathcal{F}f).$$

Let $\mathbb{H}^q_{\gamma,\mathbf{p}} = L^q(\mathbb{R}, H^\gamma_{\mathbf{p}})$, and $\mathbb{H}^q_{\gamma,\mathbf{p}}(T) = L^q((0,T), H^\gamma_{\mathbf{p}})$, and the space $\mathscr{H}^q_{\alpha,\mathbf{p}}(T)$ consists of the functions $u = u(t)$ on $[0,T]$ with values in the space of distributions on $\mathbb{R}^d$ such that $u \in \mathbb{H}^q_{\alpha,\mathbf{p}}(T)$ and $\partial_t u \in \mathbb{L}^q_{\mathbf{p}}(T)$.

Throughout this section, we always assume that

(Ha): $a(t,x)$ is uniformly continuous in $x \in \mathbb{R}^d$ locally uniformly with respect to $t \in \mathbb{R}_+$, and there exists a constant $K > 1$ such that for all $\xi \in \mathbb{R}^d$,

$$K^{-1}|\xi|^2 \leqslant |a(t,x)\xi|^2 \leqslant K|\xi|^2, \qquad \forall x \in \mathbb{R}^d.$$

The main result of this section is as follows.

Theorem 3.4. *Assume that* **(Ha)** *holds and* $\mathbf{p} \in (1,\infty)^d$ *and* $q \in (1,\infty)$. *Then for every* $f \in \mathbb{L}^q_{\mathbf{p}}(T)$, *if* $b \in \mathbb{L}^{\tilde q}_{\tilde{\mathbf{p}}}(T)$ *with* $\tilde{\mathbf{p}}, \tilde q$ *satisfying* $2/\tilde q + 1/\tilde p_1 + \cdots + 1/\tilde p_d < 1$ *and* $\tilde p_i \in [p_i,\infty)$, $\tilde q \in [q,\infty)$ *for* $1 \leqslant i \leqslant d$, *there exists a unique solution* $u \in \mathscr{W}^q_{2,\mathbf{p}}(T)$ *to the equation ((3.12)). Moreover, we have the following estimates:*

(i) there exists a constant $C_1 = C(d,\mathbf{p},q) > 0$ *such that*

$$\|\partial_t u\|_{L^q_{\mathbf{p}}(T)} + \|u\|_{\mathbb{W}^q_{2,\mathbf{p}}(T)} \leqslant C_1 \|f\|_{L^q_{\mathbf{p}}(T)}; \tag{3.14}$$

(ii) there is a constant $C_T = C(d,\mathbf{p},q,T)$ *satisfying* $\lim_{T\to 0} C_T = 0$ *such that*

$$\|u\|_{\mathbb{L}^\infty(T)} \leqslant C_T \|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \textit{if} \quad 2/q + 1/p_1 + \cdots + 1/p_d < 2, \tag{3.15}$$

and

$$\|\nabla u\|_{\mathbb{L}^\infty(T)} \leqslant C_T \|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \textit{if} \quad 2/q + 1/p_1 + \cdots + 1/p_d < 1. \tag{3.16}$$

We shall provide the proof of above result in the following subsections.

3.2.1 Constant diffusion coefficients.

Let us first assume that $a(t,x) \equiv a(t) \geqslant \mathbb{I}_{d\times d}$ is independent of x variable and $b(t,x) \equiv 0$, where I is the unit $d \times d$ matrix, i.e., consider the following PDE:

$$\partial_t u(t,x) - a_{ij}(t)\partial_{ij} u(t,x) - f(t,x) = 0, \quad u(0,x) = 0. \tag{3.17}$$

Below, a function $f(t,x)$ defined on $\mathbb{R}^{d+1}_T$ will always be extended to the whole space $\mathbb{R}^{d+1}$ automatically by setting $f(t,x) \equiv 0$ when $t < 0$ or $t > T$. Denote by $\hat{f}(\xi_0, \xi)$ the Fourier transform of f with respect to the variables t and x, i.e.,

$$\hat{f}(\xi_0,\xi) = \frac{1}{(2\pi)^{\frac{d+1}{2}}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} \mathrm{d}x \int_{-\infty}^{\infty} e^{-it\xi_0} f(t,x) \mathrm{d}t.$$

The following result extends [35, Theorem 2.1] to the general mixed-norm cases. The key tool that we use here is the Hörmander-Mihlin theorem for mixed-norm spaces. We give the detailed proof for completeness.

Lemma 3.5. *Let $\mathbf{p} \in (1,\infty)^d$ and $q \in (1,\infty)$. Then for any $f \in \mathbb{L}^q_{\mathbf{p}}(T)$, there exists a unique strong solution $u \in \mathscr{W}^q_{2,\mathbf{p}}(T)$ to equation (3.17). Moreover, estimates (3.14)-(3.16) hold true.*

Proof. (i) It suffices to prove the conclusions when $a^{ij} = \mathbb{I}_{d\times d}$, the unit matrix in $\mathbb{R}^d$. Then the general case follows by [35, Theorem 2.2], which says that whatever estimate is true for the heat equation in translation invariant spaces is also true with the same constant for (3.17) with the coefficients depending only on t provided $(a^{ij}(t)) \geqslant \mathbb{I}_{d\times d}$. In this case, it is well known that the solution u admits the following representation:

$$u(t,x) = \int_0^t \int_{\mathbb{R}^d} g(t-s, x-y) f(s,y) \mathrm{d}y\mathrm{d}s,$$

where g is the Gaussian function given by $g(t,x) = 2\pi t^{-d/2} e^{-\frac{x^2}{2t}}$. Define the operator $\mathcal{A}$ by

$$\mathcal{A}f := \nabla_x^2 u(t,x) := \int_t^T \int_{\mathbb{R}^d} \nabla_x^2 g(s-t, x-y) f(s,y) \mathrm{d}y\mathrm{d}s.$$

Then we have

$$\widehat{\mathcal{A}f}(\xi_0,\xi) = -\frac{\xi^2}{i\xi_0 + \xi^2} \hat{f}(\xi_0,\xi).$$

One can check that the function $m(\xi_0,\xi) = \frac{-\xi^2}{i\xi_0+\xi^2}$ satisfies the condition

$$\left| \xi_{i_1} \cdots \xi_{i_k} \frac{\partial^k m}{\partial \xi_{i_1} \cdots \partial \xi_{i_k}} \right| \leqslant C$$

for $k = 1, \cdots, d+1$, $i_j = 0, \cdots, d$ and $i_m \neq i_j (m \neq j)$, where $C > 0$ is a constant. Thus, according to [3, Theorem 7], $\mathcal{A}$ is a Fourier multiplier mapping from the mixed-norm Lebesgue space $\mathbb{L}^q_{\mathbf{p}}$ into itself, which in turn yields (3.14).

(ii) Let $\mathbf{p}' = (p'_1, \cdots, p'_d)$ be the conjugate exponent to $\mathbf{p}$, i.e., $1/q + 1/q' = 1$ and $1/p_i + 1/p'_i = 1$, for $i = 1, \cdots, d$. By simple computation, we have

$$\|g(t,\cdot)\|_{L^{\mathbf{p}'}(\mathbb{R}^d)} \leqslant Ct^{\frac{1}{2p'_1}+\cdots+\frac{1}{2p'_d}-\frac{d}{2}},$$

and

$$\|\nabla g(t,\cdot)\|_{L^{\mathbf{p}'}(\mathbb{R}^d)} \leqslant Ct^{\frac{1}{2p'_1}+\cdots+\frac{1}{2p'_d}-\frac{d}{2}-\frac{d}{2}} = Ct^{\frac{1}{2p'_1}+\cdots+\frac{1}{2p'_d}-d}.$$

Then, by applying Hölder's inequality for mixed-norm space (see [3, Lemma 2]), we find that for any $\mathbf{p} \in (1, \infty)^d$, $q \in (1, \infty)$ satisfying

$$\frac{1}{2p_1} + \cdots + \frac{1}{2p_d} + \frac{1}{q} < 1$$

it holds

$$|u(t,x)| \leqslant C_2 t^{1-\frac{1}{q}-(\frac{1}{2p_1}+\cdots+\frac{1}{2p_d})}\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)},$$

and for $\mathbf{p} \in (1, \infty)^d$, $q \in (1, \infty)$ satisfying

$$\frac{1}{2p_1} + \cdots + \frac{1}{2p_d} + \frac{1}{q} < \frac{1}{2},$$

it holds

$$|\nabla u(t,x)| \leqslant C_3 t^{\frac{1}{2}-\frac{1}{q}-(\frac{1}{2p_1}+\cdots+\frac{1}{2p_d})}\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)},$$

therefore we get (3.15) and (3.16). The proof is finished. □

3.2.2 Variable diffusion coefficients.

In this subsection, we consider PDE (3.12) with $b \equiv 0$, i.e.,

$$\partial_t u(t,x) - \mathscr{L}_2^a u(t,x) - f(t,x) = 0, \quad u(s,x) = \phi(x), \tag{3.18}$$

where $s \leqslant t$ and $\phi \in C_0^\infty(\mathbb{R}^d)$. We shall assume that a is smooth enough, i.e., a satisfies **(Ha)** and for all $m \in \mathbb{N}$,

$$\|\nabla^m a^{ij}(t,\cdot)\|_\infty < \infty.$$

Motivated by [67], we also need to consider the dual equation for (3.18) as following:

$$\partial_s w(s,x) + \partial_{ij}\big((a^{ij}(s,x)w(s,x)\big) + f(s,x) = 0, \quad w(t,x) = \psi(x), \tag{3.19}$$

where $s \geqslant t$ and $\psi \in C_0^\infty(\mathbb{R}^d)$. For given ϕ, $\psi \in C_0^\infty(\mathbb{R}^d)$ and $s \geqslant t$, let $u(t)$ and $w(s)$ be the unique solution of (3.18) and (3.19) respectively. We shall simply write

$$\mathcal{T}_{t,s}\phi := u(t), \quad \mathcal{T}^*_{t,s}\psi := w(s).$$

In other words, we have

$$\partial_t \mathcal{T}_{t,s}\phi = a^{ij}\partial_{ij}\mathcal{T}_{t,s}\phi, \quad \partial_s \mathcal{T}^*_{s,t}\psi = -\partial_{ij}(a^{ij}\mathcal{T}^*_{t,s}\psi).$$

By the chain rule and above equations, it is easy to see that

$$\langle \mathcal{T}_{t,s}\phi, \psi\rangle - \langle \phi, \mathcal{T}^*_{t,s}\psi\rangle = \int_s^t dr \langle \mathcal{T}_{r,s}\phi, \mathcal{T}^*_{t,r}\psi\rangle = 0.$$

That is to say

$$\langle \mathcal{T}_{t,s}\phi, \psi\rangle = \langle \phi, \mathcal{T}^*_{t,s}\psi\rangle. \tag{3.20}$$

Fix $T > 0$. For $f \in L^\infty([0,T], \mathcal{C}_0^\infty(\mathbb{R}^d))$, define

$$u(t,x) := \int_t^T \mathcal{T}_{s,T} f(s,x) ds, \quad w(s,x) := \int_0^s \mathcal{T}^*_{0,t} f(t,x) dt.$$

Then u solves the following forward equation

$$\partial_t u = a^{ij}\partial_{ij} u + f, \quad u(t)|_{t \leqslant 0} = 0, \tag{3.21}$$

and w solves the following backward equation

$$\partial_s w = -\partial_{ij}(a^{ij} w) - f, \quad w(s)|_{s \geqslant T} = 0. \tag{3.22}$$

We proceed to show the following a priori estimates.

Lemma 3.6. *For any $\mathbf{p} \in (1,\infty)^d$ and $q \in (1,\infty)$, there is a constant $C > 0$ only depending on $d, \mathbf{p}, q$ and the continuity modulus of a such that for every $f \in \mathcal{C}_0^\infty([0,T] \times \mathbb{R}^d)$,*

$$\|\nabla^2 u\|_{\mathbb{L}^q_{\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \|w\|_{\mathbb{L}^q_{\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{H}^q_{-2,\mathbf{p}}(T)}, \tag{3.23}$$

where u and w are solutions of (3.18) and (3.19) respectively. Moreover, for any $\alpha \in [0, 2 - \frac{2}{q})$,

$$\|u\|_{\mathbb{H}^\infty_{\alpha,\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \|w\|_{\mathbb{H}^\infty_{\alpha-2,\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{H}^q_{-2,\mathbf{p}}(T)}. \tag{3.24}$$

Before giving the proof of the above theorem, following the argument in [36, Lemma 1.6] we first give the following freezing lemma and Sobolev embedding theorem in Mixed-norm Lebesgue space (see proof in Appendix A.6) for later use.

Lemma 3.7. *([76, Lemma 4.1]) Let ϕ be a nonzero smooth function with compact support. Define $\phi_z(x) := \phi(x - z)$. For any $\alpha \in \mathbb{R}$ and $p \in (1,\infty)$, there exists a constant $C \geqslant 1$ depending only on α, p, ϕ such that for all $f \in H^{\alpha,p}$,*

$$\frac{1}{C}\|f\|_{H^\alpha_p} \leqslant \Big(\int_{\mathbb{R}^d} \|\phi_z f\|^p_{H^\alpha_p} dz\Big)^{1/p} \leqslant C\|f\|_{H^\alpha_p}.$$

Lemma 3.8. *For* $\mathbf{p} \in [1,\infty)^d$,

$$|||f|||_{\mathbb{L}^\infty(T)} \leqslant C|||f|||_{\mathbb{H}^\infty_{\alpha,\mathbf{p}}(T)}, \quad \alpha > \frac{1}{p_1} + \cdots + \frac{1}{p_d} \tag{3.25}$$

where $|||f|||_{\mathbb{H}^q_{\alpha,p}(T)} := \sup_{z\in\mathbb{R}^d} \|\chi^z_r f\|_{\mathbb{H}^q_{\alpha,p}(T)}$, $|||f|||_{\mathbb{L}^q_p(T)} := \sup_{z\in\mathbb{R}^d} \|\chi^z_r f\|_{\mathbb{L}^q_p(T)}$. *Here* $\chi \in \mathcal{C}^\infty_c(\mathbb{R}^d)$ *with* $\chi(x) = 1$ *for* $|x| \leqslant 1$ *and* $\chi(x) = 0$ *for* $|x| > 2$, $\chi_r(x) := r^{-d}\chi(x/r)$, $\chi^z_r(x) := \chi_r(x-z)$ *for* $r > 0$ *and* $z \in \mathbb{R}^d$.

Lemma 3.9. *([36, Lemma 1.5]) Let* $a : \mathbb{R} \to \mathbb{R}^d \otimes \mathbb{R}^d$ *be a measurable and symmetric matrix-valued function and there exists a constant* $\delta \geqslant 1$ *such that for all* $t \in [0,\infty)$

$$\delta^{-1}|\xi|^2 \leqslant a^{ij}(t)\xi_i\xi_j \leqslant \delta|\xi|^2, \quad \forall \xi \in \mathbb{R}^d.$$

Let $T \leqslant \infty$, $p \in (1,\infty)$, *and let* $u \in L^p((0,T)\times\mathbb{R}^d) = \mathbb{L}^p_p(T)$ *be a solution of the equation*

$$\partial_t u = a^{ij}\partial_{ij}u + f, \quad u(0,x) = 0$$

with $f \in \mathbb{L}^p_p(T)$. *Then*

$$\|u\|_{\mathbb{L}^p_p(T)} \leqslant N(d,p)\|f\|_{\mathbb{L}^p_p(T)}.$$

Lemma 3.10. *Let* $T \in [0,\infty)$, $p \in (1,\infty)$, $n \in \mathbb{N}$. *For* $k = 1, \cdots, n$, *let* $a_k : \mathbb{R} \to \mathbb{R}^d \otimes \mathbb{R}^d$ *be measurable and there exists a constant* $\delta \geqslant 1$ *such that for all* $t \in [0,\infty)$

$$\delta^{-1}|\xi|^2 \leqslant a^{ij}_k(t)\xi_i\xi_j \leqslant \delta|\xi|^2, \quad \forall \xi \in \mathbb{R}^d.$$

Let $\lambda_k \in (0,\infty)$, $\gamma_k \in \mathbb{R}$, *and* $u_k \in \mathbb{H}^p_{\gamma_k+2,p}(T)$ *be the solution to the equation*

$$\partial_t u^k = a^{ij}_k \partial_{ij}u^k + f^k, \quad u^k(0,x) = 0$$

with $f \in \mathbb{H}^p_{\gamma_k,p}(T)$. *Denote* $\Lambda_k = (\lambda_k - \Delta)^{\gamma_k/2}$, *then for* $i = 2, ..., d$, *we have*

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{k=1}^n \|\Lambda_k \Delta u^k(t,\cdot,x_i,\cdots,x_d)\|^p_{L^p(\mathbb{R}^{i-1})} dx_i \cdots dx_d dt \\
&\quad \leqslant N \sum_{k=1}^n \int_0^T \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \|\Lambda_k f^k(t,\cdot,x_i,\cdots,x_d)\|^p_{L^p(\mathbb{R}^{i-1})} \\
&\qquad\qquad \prod_{j\neq k} \|\Lambda_j \Delta u^j(t,\cdot,x_i,\cdots,x_d)\|^p_{L^p(\mathbb{R}^{i-1})} dx_i \cdots dx_d dt.
\end{aligned} \tag{3.26}$$

Proof. Without loss of generality we may assume $\gamma_k = 0$. Define $v^k = \Delta u^k$. For fixed $i = \{2, ..., d\}$, take $X = (x^1, \cdots, x^n) \in \mathbb{R}^{nd}$ with $x^j \in \mathbb{R}^d$ and $x^j_i = x_i \in \mathbb{R}$, $x^j_{i+1} = x_{i+1} \in \mathbb{R}$, $\cdots$, $x^j_d = x_d \in \mathbb{R}$ for $1 \leqslant j \leqslant n$, hence actually $X \in \mathbb{R}^{d+(n-1)(i-1)}$. For such X, we define

$$V(t,X) = v^1(t,x^1) \cdot ... \cdot v^n(t,x^n).$$

then by the fact that a is independent of space variable and

$$\partial_t v^k(t,x^k) = a_k^{ij}\partial_{ij}\Delta_{x^k}u^k(t,x^k) + \Delta_{x^k}f^k(t,x^k), \quad k=1,\cdots,n,$$

we get

$$\partial_t V(t,X) = \mathbb{P}V(t,X) + F(t,X),$$

where

$$\mathbb{P}V = a_k^{ij}\frac{\partial^2 V}{\partial x_i^k \partial x_j^k},$$

$$F(t,X) = \Delta_{x^k}G^k(t,X), \quad G^k(t,X) = f^k(t,x^k)\prod_{j\neq k} v^j(t,x^j).$$

By classical result, i.e. Lemma 3.9, we have

$$\|V\|_{L^p((0,T)\times\mathbb{R}^{d+(n-1)(i-1)})} \leqslant N\sum_j \|G^j\|_{L^p((0,T)\times\mathbb{R}^{d+(n-1)(i-1)})},$$

which is exactly (3.26). The lemma is proved. □

According to [67], we have the following estimate about equations (3.21) and (3.31).

Lemma 3.11. *([67, Theorem 3.2]) For any* $\mathbf{p}\in(1,\infty)^d$ *and* $q\in(1,\infty)$*, there is a constant* $C>0$ *depending only on* $d,\mathbf{p},q,T$ *and the continuity modulus of* a *such that for every* $f\in C_0^\infty([0,T]\times\mathbb{R}^d)$*,*

$$\|\nabla^2 u\|_{\mathbb{L}^q_{\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \|w\|_{\mathbb{L}^q_{\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{H}^q_{-2,\mathbf{p}}(T)}, \tag{3.27}$$

where u *and* w *are solutions of (3.18) and (3.19) respectively. Moreover, for any* $\alpha\in[0,2-\frac{2}{q})$*, we have*

$$\|u\|_{\mathbb{H}^\infty_{\alpha,\mathbf{p}}(T)} \leqslant C_T\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \|w\|_{\mathbb{H}^\infty_{\alpha-2,\mathbf{p}}(T)} \leqslant C_T\|f\|_{\mathbb{H}^q_{-2,\mathbf{p}}(T)}. \tag{3.28}$$

where $C_T>0$ *is a constant satisfying* $\lim_{T\to 0} C_T = 0$*.*

With the above preparation, we can give:

Proof of Lemma 3.6. Let $\mathbf{p}=(p_1,p_2,\cdots,p_d)\in(1,\infty)^d$ and $q\in(1,\infty)$. We divide the proof into five steps: we first prove estimate (3.27) in step 1-4, and in the fifth step we show estimate (3.28).

Step 1. [Case $p_1=\cdots=p_d\in(1,\infty)$ and $q\in(1,\infty)$]. In this case, the estimate (3.27) was proved by [67, Theorem 3.3].

Step 2. [Case $p_1=\cdots=p_{d-1}\in(1,\infty)$ and $p_d=q\in(1,\infty)$]. We only prove the estimate for w since the estimate for u is similar and easier. By duality and the same argument as in the proof of [67, Theorem 3.3], it is sufficient to prove the desired estimate when

$q = p_d = np_{d-1} = \cdots = np_1 =: np$ for $n \in \mathbb{N}_+$ and $p \in (1, \infty)$. That is to say, we shall prove:

$$\|w\|_{L^{np}([0,T]\times\mathbb{R}, L^p(\mathbb{R}^{d-1}))} \leqslant C\|f\|^{np}_{\mathbb{H}^{np}_{-2,\mathbf{p}}(T)}, \quad \mathbf{p} = (p, \cdots, p, np).$$

Take a non-negative smooth function ϕ supported in the ball $B_r := \{x \in \mathbb{R}^d : |x| < r\}$ and $\int_{\mathbb{R}^d} |\phi|^p dx = 1$, where r is a small constant which will be determined below. For $x, z \in \mathbb{R}^d$, $s \in \mathbb{R}_+$, define $\phi_z(x) := \phi(x - z)$, $w_z(s, x) := w(s, x)\phi_z(x)$, $f_z(s, x) := f(s, x)\phi_z(x)$ and $a_z(s) := a(s, z)$. Then we can write

$$\partial_t w_z + \partial_{ij}(a^{ij}_z w_z) + g_z = 0, \quad w_z(T, x) = 0, \tag{3.29}$$

where

$$g_z = f_z + \partial_{ij}(a^{ij}w)\phi_z - \partial_{ij}(a^{ij}_z w\phi_z).$$

Below, for any $\gamma \in \mathbb{R}$ and fixed $x_d \in \mathbb{R}$, we denote by $\|f(\cdot, x_d)\|_{H^\gamma_p(\mathbb{R}^{d-1})} := \|((1 - \Delta)^{\gamma/2} f)(\cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}$, and drop the time variable for simplicity. Notice that

$$g_z = f\phi_z - 2\partial_j(a^{ij}w)\partial_i\phi_z - a^{ij}w\partial_i\partial_j\phi_z + \partial_i\partial_j((a^{ij} - a^{ij}_z)w_z).$$

By the continuity of a, we have

$$\begin{aligned}\Big(\int_{\mathbb{R}^d} \|g_z(\cdot, x_d)\|^p_{H^{-2}_p(\mathbb{R}^{d-1})} \mathrm{d}z\Big)^{1/p} \leqslant{}& C\|f(\cdot, x_d)\|_{H^{-2}_p(\mathbb{R}^{d-1})} \\ &+ C_r \sum_{i,j} \|(a^{ij}w)(\cdot, x_d)\|_{H^{-1}_p(\mathbb{R}^{d-1})} \\ &+ C_r \sum_{i,j} \|a^{ij}w(\cdot, x_d)\|_{H^{-2}_p(\mathbb{R}^{d-1})} + c_r\|w(\cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})},\end{aligned}$$

where $C_r > 0$ and $\lim_{r\to 0} c_r = 0$. Let ρ_n be a family of standard mollifiers and $a_n(t, x) := a(t, \cdot) * \rho_n(x)$ be the mollifying approximation of a. For every $\varepsilon > 0$, we can take n large enough such that

$$\begin{aligned}\sum_{i,j} \|(a^{ij}w)(\cdot, x_d)\|_{H^{-1}_p(\mathbb{R}^{d-1})} + \sum_{i,j} \|a^{ij}w(\cdot, x_d)\|_{H^{-2}_p(\mathbb{R}^{d-1})} \\ \leqslant{}& C\|(aw)(\cdot, x_d)\|_{H^{-1}_p(\mathbb{R}^{d-1})} \\ \leqslant{}& C\|(a_n w)(\cdot, x_d)\|_{H^{-1}_p(\mathbb{R}^{d-1})} + C\|((a - a_n)w)(\cdot, x_d)\|_{H^{-1}_p(\mathbb{R}^{d-1})} \\ \leqslant{}& C_n\|w(\cdot, x_d)\|_{H^{-1}_p(\mathbb{R}^{d-1})} + c_{1/n}\|w(\cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})} \\ \leqslant{}& C_n\|w(\cdot, x_d)\|_{H^{-2}_p(\mathbb{R}^{d-1})} + \varepsilon\|w(\cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})},\end{aligned}$$

where the last step is due to the interpolation and Young's inequalities. Hence, we get

$$\begin{aligned}\Big(\int_{\mathbb{R}^d} \|g_z(\cdot, x_d)\|^p_{H^{-2}_p(\mathbb{R}^{d-1})} \mathrm{d}z\Big)^{1/p} \leqslant{}& C\|f(\cdot, x_d)\|_{H^{-2}_p(\mathbb{R}^{d-1})} \\ &+ C_r\|w(\cdot, x_d)\|_{H^{-2}_p(\mathbb{R}^{d-1})} + c_r\|w(\cdot, x_d)\|_{L^p(\mathbb{R}^{d-1})}.\end{aligned} \tag{3.30}$$

Observe that

$$\begin{aligned}\|w\|^{np}_{L^{np}([0,T]\times\mathbb{R},L^p(\mathbb{R}^{d-1}))} &= \int_0^T\int_{\mathbb{R}} \Big(\int_{\mathbb{R}^d} \|w(t,\cdot,x_d)\phi_z\|^p_{L^p(\mathbb{R}^{d-1})} dz \Big)^n \mathrm{d}x_d\mathrm{d}t \\ &= \int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}^{nd}} \prod_{k=1}^n \|w_{z_k}(t,\cdot,x_d)\|^p_{L^p(\mathbb{R}^{d-1})}\mathrm{d}z_1\cdots\mathrm{d}z_n\mathrm{d}x_d\mathrm{d}t. \end{aligned} \tag{3.31}$$

Using Lemma 3.10, we can deduce that

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}} \prod_{k=1}^n \|w_{z_k}(t,\cdot,x_d)\|^p_{L^p(\mathbb{R}^{d-1})}\mathrm{d}x_d\mathrm{d}t \\ &\qquad \leqslant N\sum_{k=1}^n \int_0^T\int_{\mathbb{R}} \|g_{z_k}(t,\cdot,x_d)\|^p_{H_p^{-2}(\mathbb{R}^{d-1})} \prod_{l\neq k}\|w_{z_l}(t,\cdot,x_d)\|^p_{L^p(\mathbb{R}^{d-1})}\mathrm{d}x_d\mathrm{d}t,\end{aligned}$$

which together with (3.30) and (3.31) implies

$$\begin{aligned}\|w\|^{np}_{L^{np}([0,T]\times\mathbb{R},L^p(\mathbb{R}^{d-1}))} &\leqslant C_0\sum_{k=1}^n\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}^{nd}} \|g_{z_k}(t,\cdot,x_d)\|^p_{H_p^{-2}(\mathbb{R}^{d-1})} \\ &\qquad\qquad \times\prod_{l\neq k}\|w_{z_l}(t,\cdot,x_d)\|^p_{L^p(\mathbb{R}^{d-1})}\mathrm{d}z_1\cdots\mathrm{d}z_n\mathrm{d}x_d\mathrm{d}t \\ &= C_0 n\int_0^T\int_{\mathbb{R}}\Big(\int_{\mathbb{R}^d}\|g_z(t,\cdot,x_d)\|^p_{H_p^{-2}(\mathbb{R}^{d-1})}\mathrm{d}z\Big) \\ &\qquad\qquad \times\Big(\int_{\mathbb{R}^d}\|w_z(t,\cdot,x_d)\|^p_{L^p(\mathbb{R}^{d-1})}\mathrm{d}z\Big)^{n-1}\mathrm{d}x_d\mathrm{d}t \\ &\leqslant C\int_0^T\int_{\mathbb{R}}\Big(\int_{\mathbb{R}^d}\|g_z(t,\cdot,x_d)\|^p_{H_p^{-2}(\mathbb{R}^{d-1})}dz\Big) \\ &\qquad\qquad \times\|w(t,\cdot,x_d)\|^{(n-1)p}_{L^p(\mathbb{R}^{d-1})}\mathrm{d}x_d\mathrm{d}t \\ &\leqslant C\int_0^T\int_{\mathbb{R}}\|f(t,\cdot,x_d)\|^{np}_{H_p^{-2}(\mathbb{R}^{d-1})}\mathrm{d}x_d\mathrm{d}t \\ &\qquad\qquad + C_r\int_0^T\int_{\mathbb{R}}\|w(t,\cdot,x_d)\|^{np}_{H_p^{-2}(\mathbb{R}^{d-1})}\mathrm{d}x_d\mathrm{d}t \\ &\qquad\qquad + c_r\|w\|^{np}_{L^{np}([0,T]\times\mathbb{R},L^p(\mathbb{R}^{d-1}))},\end{aligned}$$

where the last inequality follows from Hölder's inequality and Young's inequality for product. Let r be small enough so that $c_r < 1$, we can get that

$$\begin{aligned}\|w\|^{np}_{L^{np}([0,T]\times\mathbb{R},L^p(\mathbb{R}^{d-1}))} &\leqslant C\Big(\int_0^T\int_{\mathbb{R}}\|f(t,\cdot,x_d)\|^{np}_{H_p^{-2}(\mathbb{R}^{d-1})}\mathrm{d}x_d\mathrm{d}t \\ &\qquad + \int_0^T\int_{\mathbb{R}}\|w(t,\cdot,x_d)\|^{np}_{H_p^{-2}(\mathbb{R}^{d-1})}\mathrm{d}x_d\mathrm{d}t\Big).\end{aligned} \tag{3.32}$$

It remains to control the last term on the right hand side of the above inequality. To this end, let $\kappa^z_{s,t} := \int_s^t a_z(u)\mathrm{d}u$ and

$$P^z_{s,t}(x, x-y) := \frac{1}{\sqrt{(2\pi)^d \det(\kappa^z_{s,t})}} e^{-\frac{(\kappa^z_{s,t})^{-1}|x-y|^2}{2(t-s)}}.$$

Then the solution of equation (3.29) is given by

$$w_z(t,x) = \int_t^T \int_{\mathbb{R}^d} P^z_{t,u}(x, x-y) g_z(u,y)\mathrm{d}y\mathrm{d}u.$$

By $(\mathbf{H}a)$ and a standard interpolation technique, we get that for any $\alpha \in [0,2)$,

$$\|w_z(t,\cdot,x_d)\|_{H_p^{\alpha-2}(\mathbb{R}^{d-1})} \leqslant C \int_t^T (u-t)^{-\frac{\alpha}{2}} \|g_z(u,\cdot,x_d)\|_{H_p^{-2}(\mathbb{R}^{d-1})}\mathrm{d}u.$$

Thus by Minkowski's inequality we have

$$\begin{aligned} \|w(t,\cdot,x_d)\|_{H_p^{\alpha-2}(\mathbb{R}^{d-1})} &\leqslant \Big(\int_{\mathbb{R}^d} \|w_z(t,\cdot,x_d)\|^p_{H_p^{\alpha-2}(\mathbb{R}^{d-1})}\mathrm{d}z \Big)^{1/p} \\ &\leqslant \int_t^T (u-t)^{-\frac{\alpha}{2}} \Big(\int_{\mathbb{R}^d} \|g_z(u,\cdot,x_d)\|^p_{H_p^{-2}(\mathbb{R}^{d-1})}\mathrm{d}z \Big)^{1/p}\mathrm{d}u. \end{aligned}$$

Using (3.30) and the similar argument as in the proof of (3.32), we further have

$$\begin{aligned} \|w(t,\cdot,x_d)\|_{H_p^{\alpha-2}(\mathbb{R}^{d-1})} \leqslant C \int_t^T (u-t)^{-\frac{\alpha}{2}} \Big(&\|f(u,\cdot,x_d)\|_{H_p^{-2}(\mathbb{R}^{d-1})} \\ &+ \|w(u,\cdot,x_d)\|_{H_p^{-2}(\mathbb{R}^{d-1})} \Big)\mathrm{d}u. \end{aligned}$$

Let $\frac{1}{q'} + \frac{1}{np} = 1$, then for any $\alpha \in [0, 2 - \frac{2}{np})$, we get by Hölder's inequality that

$$\begin{aligned} \|w(t,\cdot,x_d)\|^{np}_{H_p^{\alpha-2}(\mathbb{R}^{d-1})} &\leqslant C \Big(\int_t^T (u-t)^{-\frac{q'\alpha}{2}}\mathrm{d}u \Big)^{np/q'} \int_t^T \Big(\|f(u,\cdot,x_d)\|_{H_p^{-2}(\mathbb{R}^{d-1})} \\ &\qquad + \|w(u,\cdot,x_d)\|_{H_p^{-2}(\mathbb{R}^{d-1})} \Big)^{np}\mathrm{d}u \\ &\leqslant C_T \int_t^T \Big(\|f(u,\cdot,x_d)\|^{np}_{H_p^{-2}(\mathbb{R}^{d-1})} + \|w(u,\cdot,x_d)\|^{np}_{H_p^{-2}(\mathbb{R}^{d-1})} \Big)\mathrm{d}u, \end{aligned} \tag{3.33}$$

where $C_T > 0$ satisfying $\lim_{T\to 0} C_T = 0$. Then by taking $\alpha = 0$ and Gronwall's inequality we can obtain

$$\sup_{s\in[0,T]} \|w(s,\cdot,x_d)\|^{np}_{H_p^{-2}(\mathbb{R}^{d-1})} \leqslant C_T \int_t^T \|f(u,\cdot,x_d)\|^{np}_{H_p^{-2}(\mathbb{R}^{d-1})}\mathrm{d}u, \tag{3.34}$$

which in particular implies that

$$\|w\|^{np}_{\mathbb{H}^{\infty}_{-2,\mathbf{p}}} \leqslant C_T \|f\|^{np}_{\mathbb{H}^{np}_{-2,\mathbf{p}}}.$$

Taken this back into (3.32) yields that

$$\|w\|_{L^{np}([0,T]\times\mathbb{R},L^p(\mathbb{R}^{d-1}))} \leqslant C\|f\|^{np}_{\mathbb{H}^{np}_{-2,\mathbf{p}}(T)}, \quad \mathbf{p} = (p,\cdots,p,np).$$

Step 3. [Case $p_1 = \cdots = p_{d-j} \in (1,\infty)$ and $p_{d-j+1} = \cdots = p_d = q \in (1,\infty)$ with any $1 \leqslant j \leqslant d-1$]. This can be proved by following exactly the same arguments as in the proof of step 2, except that we need to use Lemma 3.10 (3.26) with $i = d-j+1$, we omit the details.

Step 4. [Interpolation] We develop an interpolation scheme to show the following claim:

$$\text{for every } 1 \leqslant j \leqslant d-1, (3.27) \text{ holds with } p_1 = \cdots = p_{d-j} \in (1,\infty) \\ \text{and } p_{d-j+1}, p_{d-j+2}, \cdots, p_d, q \in (1,\infty). \tag{3.35}$$

In particular, when $j = d-1$, we get the desired result.

Interpolate the results in step 1 and step 2, we can get that (3.27) holds when $p_1 = \cdots = p_{d-1} \in (1,\infty)$ and $p_d, q \in (1,\infty)$. Thus, assertion (3.35) holds for $j = 1$. Assume that (3.35) holds for some $j = n-1 \leqslant d-2$, we proceed to show that (3.35) is true for n. For this, we first interpolate $p_1 = \cdots = p_d \in (1,\infty)$ and $q \in (1,\infty)$ with $p_1 = \cdots = p_{d-j} \in (1,\infty)$ and $p_{d-j+1} = \cdots = p_d = q \in (1,\infty)$ (both of which hold according to step 3) to get that the (3.27) holds for $p_1 = \cdots = p_{d-j} \in (1,\infty)$ and $p_{d-j+1} = p_{d-j+2} = \cdots = p_d, q \in (1,\infty)$. Then we interpolate $p_1 = \cdots = p_{d-j} \in (1,\infty)$ and $p_{d-j+1} = p_{d-j+2} = \cdots = p_d, q \in (1,\infty)$ with $p_1 = \cdots = p_{d-1} \in (1,\infty)$ and $p_d, q \in (1,\infty)$ (which holds by induction assumption for $j = 1$) to get that (3.27) holds for $p_1 = \cdots = p_{d-j} \in (1,\infty)$ and $p_{d-j+1} = \cdots = p_{d-1}, p_d, q \in (1,\infty)$. Again we interpolate $p_1 = \cdots = p_{d-j} \in (1,\infty)$ and $p_{d-j+1} = \cdots = p_{d-1}, p_d, q \in (1,\infty)$ with $p_1 = \cdots = p_{d-2} \in (1,\infty)$ and $p_{d-1}, p_d, q \in (1,\infty)$ (which holds by induction assumption for $j = 2$) to get that (3.27) holds for $p_1 = \cdots = p_{d-j} \in (1,\infty)$ and $p_{d-j+1} = \cdots = p_{d-2}, p_{d-1}, p_d, q \in (1,\infty)$. Keep interpolating with induction assumption for $j = 3, \cdots, n-1$, we can get that (3.27) holds for $p_1 = \cdots = p_{d-j} \in (1,\infty)$ and $p_{d-j+1}, p_{d-j+2}, \cdots, p_d, q \in (1,\infty)$.

Step 5. Finally, we proceed to prove estimate (3.28). With the same argument as in the previous 4 steps, it is sufficient to prove the following estimate:

$$\|w\|^{np}_{\mathbb{H}^{\infty}_{\alpha-2,\mathbf{p}}} \leqslant C_T\|f\|^{np}_{\mathbb{H}^{np}_{-2,\mathbf{p}}}, \quad \mathbf{p} = (p,\cdots,p,np), \quad \alpha \in [0, 2-\frac{2}{np}),$$

where $\lim_{T\to 0} C_T = 0$. In fact, by (3.33) and (3.34), we get for any $\alpha \in [0, 2-\frac{2}{np})$,

$$\begin{aligned}\sup_{s\in[0,T]} \int_{\mathbb{R}} \|w(s,\cdot,x_d)\|^{np}_{H^{\alpha-2}_p(\mathbb{R}^{d-1})} \mathrm{d}x_d &\leqslant \int_{\mathbb{R}} \sup_{s\in[0,T]} \|w(s,\cdot,x_d)\|^{np}_{H^{\alpha-2}_p(\mathbb{R}^{d-1})} \mathrm{d}x_d \\ &\leqslant C_T\|f\|^{np}_{\mathbb{H}^{np}_{-2,\mathbf{p}}}.\end{aligned}$$

The whole proof can be finished. □

3.2.3 Proof of Theorem 3.4

Now, we can give:

Proof of Theorem 3.4. By standard continuity method, it suffices to establish the a priori estimates (3.14)-(3.16). We divide the proof into two steps.

(i) (Case $b \equiv 0$) For $T > 0$ and $p, q \in (1, \infty)$, let $u \in \mathscr{W}^q_{2,\mathbf{p}}(T)$ and $f \in \mathbb{L}^q_{\mathbf{p}}(T)$ satisfy (3.18). Let ρ_n be a family of mollifiers. Define

$$u_n(t, x) := u(t, \cdot) * \rho_n(x), \quad a_n(t, x) := a(t, \cdot) * \rho_n(x), \quad f_n(t, x) := f(t, \cdot) * \rho_n(x).$$

It is easy to see that u_n satisfies

$$\partial_t u_n = a^{ij}_n \partial_{ij} u_n + g_n, \quad u_n(0) = 0,$$

where

$$g_n := f_n + (a^{ij} \partial_{ij} u) * \rho_n - a^{ij}_n \partial_{ij} u_n.$$

Then by (3.27), we have

$$\|\nabla^2 u_n\|_{\mathbb{L}^q_{\mathbf{p}}(T)} \lesssim \|f_n\|_{\mathbb{L}^q_{\mathbf{p}}(T)} + \|(a^{ij} \partial_{ij} u) * \rho_n - a^{ij}_n \partial_{ij} u_n\|_{\mathbb{L}^q_{\mathbf{p}}(T)},$$

and by Sobolev embedding theorem (Lemma A.6) and (3.28), for $\frac{1}{p_1} + \cdots + \frac{1}{p_d} + \frac{2}{q} < 2$, for small $\epsilon \in (0, 2 - 2/q)$, we have

$$\|u_n\|_{\mathbb{L}^\infty(T)} \leqslant C\|u_n\|_{\mathbb{H}^\infty_{2-2/q-\epsilon,\mathbf{p}}(T)} \leqslant C\|f_n\|_{\mathbb{L}^q_{\mathbf{p}}(T)} + \|(a^{ij} \partial_{ij} u) * \rho_n - a^{ij}_n \partial_{ij} u_n\|_{\mathbb{L}^q_{\mathbf{p}}(T)},$$

for $\frac{1}{p_1} + \cdots + \frac{1}{p_d} + \frac{2}{q} < 1$, for small $\epsilon \in (0, 1 - 2/q)$

$$\|\nabla u_n\|_{\mathbb{L}^\infty(T)} \leqslant C\|\nabla u_n\|_{\mathbb{H}^\infty_{1-2/q-\epsilon,\mathbf{p}}(T)} \leqslant C\|f_n\|_{\mathbb{L}^q_{\mathbf{p}}(T)} + \|(a^{ij} \partial_{ij} u) * \rho_n - a^{ij}_n \partial_{ij} u_n\|_{\mathbb{L}^q_{\mathbf{p}}(T)}.$$

Letting $n \to \infty$, we obtain

$$\|\nabla^2 u\|_{\mathbb{L}^q_{\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \|u\|_{\mathbb{H}^\infty_{2-2/q-\epsilon,\mathbf{p}}(T)} \leqslant C\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \tag{3.36}$$

$$\|\nabla u\|_{\mathbb{L}^\infty(T)} \leqslant C\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \text{if} \quad \frac{1}{p_1} + \cdots + \frac{1}{p_d} + \frac{2}{q} < 1, \tag{3.37}$$

$$\|u\|_{\mathbb{L}^\infty(T)} \leqslant C\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \text{if} \quad \frac{1}{p_1} + \cdots + \frac{1}{p_d} + \frac{2}{q} < 2. \tag{3.38}$$

(ii) Let $\frac{1}{p_i} = \frac{1}{\tilde{p}_i} + \frac{1}{\hat{p}_i}$ and $\frac{1}{q} = \frac{1}{\tilde{q}} + \frac{1}{\hat{q}}$, by Hölder's inequality and Sobolev embedding theory (Lemma A.6), we get

$$\begin{aligned} \|b \cdot \nabla u\|_{\mathbb{L}^q_{\mathbf{p}}(T)} \leqslant C\|b\|_{\mathbb{L}^{\tilde{q}}_{\tilde{\mathbf{p}}}(T)} \|\nabla u\|_{\mathbb{L}^{\hat{q}}_{\hat{\mathbf{p}}}(T)} &\leqslant C_T\|b\|_{\mathbb{L}^{\tilde{q}}_{\tilde{\mathbf{p}}}(T)} \|u\|_{\mathbb{H}^{\hat{q}}_{1+\theta,\mathbf{p}}(T)} \\ &\leqslant C_T\|b\|_{\mathbb{L}^{\tilde{q}}_{\tilde{\mathbf{p}}}(T)} \|u\|_{\mathbb{H}^\infty_{1+\theta,\mathbf{p}}(T)}. \end{aligned} \tag{3.39}$$

where $\theta \in (\frac{1}{\tilde p_1}+\cdots+\frac{1}{\tilde p_d}, 1-\frac{2}{q}) \subset (\frac{1}{\tilde p_1}+\cdots+\frac{1}{\tilde p_d}, 1-\frac{2}{\tilde q})$. We have by the result of the first step (3.36) that

$$\|u\|_{\mathbb{H}^\infty_{1+\theta,\mathbf{p}}(T)} \leqslant C_T\Big(\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)} + \|b\cdot\nabla u\|_{\mathbb{L}^q_{\mathbf{p}}(T)}\Big) \leqslant C_T\Big(\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)} + \|b\|_{\mathbb{L}^{\tilde q}_{\tilde{\mathbf{p}}}(T)}\|u\|_{\mathbb{H}^\infty_{1+\theta,\mathbf{p}}(T)}\Big),$$

By choosing T small enough so that $C_T\|b\|_{\mathbb{L}^{\tilde q}_{\tilde{\mathbf{p}}}(T)} < 1$, we have

$$\|u\|_{\mathbb{H}^\infty_{1+\theta,\mathbf{p}}(T)} \leqslant C(d,\mathbf{p},q)\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}.$$

Then

$$\begin{aligned}\|\nabla^2 u\|_{\mathbb{L}^q_{\mathbf{p}}(T)} &\leqslant C_T\Big(\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)} + \|b\cdot\nabla u\|_{\mathbb{L}^q_{\mathbf{p}}(T)}\Big) \leqslant C_T\Big(\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)} + \|b\|_{\mathbb{L}^{\tilde q}_{\tilde{\mathbf{p}}}(T)}\|u\|_{\mathbb{H}^\infty_{1+\theta,\mathbf{p}}(T)}\Big)\\ &\leqslant C(d,\mathbf{p},q)\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}.\end{aligned}$$

Which implies

$$\|\partial_t u\|_{\mathbb{L}^q_{\mathbf{p}}(T)} + \|u\|_{\mathbb{W}^q_{2,\mathbf{p}}(T)} \leqslant C(d,\mathbf{p},q)\|f\|_{\mathbb{L}^q_{\mathbf{p}}(T)}$$

With the similar argument, combining (3.38) and (3.39) we get estimates (3.15) hold, combing (3.37) and (3.39) we get (3.16). ☐

3.3 Krylov estimates and existence of weak solutions

We first give the existence result for weak solutions and Krylov's estimate, which will play an important role below.

Theorem 3.12. *Assume* **(H***a***)** *holds and* $b \in \mathbb{L}^q_{\mathbf{p}}(T)$ $q, p_1, \cdots, p_d \in (2,\infty)$ *and* $2/q + 1/p_1 + \cdots + 1/p_d < 1$. *Then there exists a weak solution* $(X_t)_{t\geqslant 0}$ *to solution to SDE (3.3). Moreover, for any non-negative function* $f \in \mathbb{L}^{\hat q}_{\hat{\mathbf{p}}}(T)$ *with* $\hat q, \hat p_1, \cdots, \hat p_d \in (1,\infty)$ *and* $2/\hat q + 1/\hat p_1 + \cdots + 1/\hat p_d < 2$, *we have*

$$\mathbb{E}\left(\int_0^t |f(s,X_s)|\mathrm{d}s\right) \leqslant C\|f\|_{\mathbb{L}^{\hat q}_{\hat{\mathbf{p}}}(T)}, \tag{3.40}$$

where $C = C(d,\hat{\mathbf{p}},\hat q,\|b\|_{\mathbb{L}^q_{\mathbf{p}}(T)})$ *is a positive constant.*

Proof. Firstly, since we already establish the estimates (3.14), (3.15) and (3.16), by following the same argument as in [73, Theorem 2.1], we can show that (3.40) holds when $b \equiv 0$. More precisely, for any $0 < S < T$ and non-negative function $f \in \mathbb{L}^{\hat q}_{\hat{\mathbf{p}}}(S,T)$ with $2/\hat q + 1/\hat p_1 + \cdots + 1/\hat p_d < 2$, there exists a constant $C(d,\mathbf{p},q) > 0$ such that

$$\mathbb{E}\left(\int_S^T |f(t,Y_t)|\mathrm{d}t\right) \leqslant C\|f\|_{\mathbb{L}^{\hat q}_{\hat{\mathbf{p}}}(S,T)}, \tag{3.41}$$

where Y_t solves the following SDE without drift

$$\mathrm{d}Y_t = \sigma(t,Y_t)\mathrm{d}W_t, \quad Y_0 = x.$$

In order to make our thesis self-complete, we give more details of getting (3.41). By Theorem 3.4, there exists a unique solution u to the following backward equation on $[0,T]$

$$\partial_t u + a^{ij}\partial_{ij} u = f, \quad u(T,x) = 0$$

for $f \in \mathbb{L}^{\hat{q}}_{\hat{\mathbf{p}}}(T)$ with $\hat{q}, \hat{p}_1, \cdots, \hat{p}_d \in (1,\infty)$ and $2/\hat{q} + 1/\hat{p}_1 + \cdots + 1/\hat{p}_d < 2$. Furthermore by (3.15) we have

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |u(t,x)| \leqslant \|f\|_{\mathbb{L}^{\hat{q}}_{\hat{\mathbf{p}}}(T)}.$$

Along the same lines of the proof of Theorem 3.4, the above statement also holds for $f \in \mathbb{L}^{d+1}_{d+1}(T) \cap \mathbb{L}^{\hat{q}}_{\hat{\mathbf{p}}}(T)$. Since $\mathbb{L}^{d+1}_{d+1}(T) \cap \mathbb{L}^{\hat{q}}_{\hat{\mathbf{p}}}(T)$ is dense in $\mathbb{L}^{\hat{q}}_{\hat{\mathbf{p}}}(T)$, so its enough to prove (3.41) holds for $f \in \mathbb{L}^{d+1}_{d+1}(T) \cap \mathbb{L}^{\hat{q}}_{\hat{\mathbf{p}}}(T)$. We take a nonnegative smooth function ρ defined on $\mathbb{R}^{d+1}$ with support in $\{x \in \mathbb{R}^{d+1} : |x| < 1\}$ and $\int_{\mathbb{R}^{d+1}} \rho(t,x)dtdx = 1$. Set $\rho_n(t,x) := n^{d+1}\rho(nt,nx)$ and extend $u(s)$ on $s \in \mathbb{R}$ by setting $u(s,\cdot) = 0$ for $s \geqslant T$ and $u(s,\cdot) = u(0,\cdot)$ for $s \leqslant 0$. Define

$$u_n(t,x) := u \star \rho_n(t,x)$$

and

$$f_n := \partial_t u_n - a^{ij}\partial_{ij} u_n.$$

Then we have

$$\begin{aligned}\|f - f_n\|_{\mathbb{L}^{d+1}_{d+1}(T)} &\leqslant \|\partial_t(u_n - u)\|_{\mathbb{L}_{\mathbb{L}^{d+1}_{d+1}(T)}} + K\|\partial_{ij}(u_n - u)\|_{\mathbb{L}_{\mathbb{L}^{d+1}_{d+1}(T)}} \\ &\leqslant \|\partial_t(u_n - u)\|\mathbb{L}^{d+1}_{d+1}(T) + K\|u_n - u\|_{\mathbb{H}^{d+1}_{2,d+1}(T)} \to 0 \text{ as } n \to \infty.\end{aligned}$$

So, by classical Krylov's estiamte (cf. [34, Lemma 5.1]), we have

$$\lim_{n\to\infty} E\Big(\int_0^T |f_n(s,Y_s) - f(s,Y_s)|ds\Big) \leqslant \lim_{n\to\infty} \|f_n - f\|_{\mathbb{L}^{d+1}_{d+1}(T)} = 0. \tag{3.42}$$

Now we use Itô's formula,

$$u_n(t,Y_t) = u_n(0,Y_0) + \int_0^t f_n(s,Y_s)ds + \int_0^t \partial_i u_n(s,Y_s)\sigma^{ik}(s,Y_s)dW^k_s, \quad \forall t \in [0,T].$$

In view of

$$\sup_{(s,x)\in[0,\infty)\times\mathbb{R}^d} |\partial_i u_n(s,x)| \leqslant C_n,$$

by Doob's optional theorem, we have

$$E\Big[\int_S^T \partial_i u_n(s,Y_s)\sigma^{ik}(s,Y_s)dW^k_s\Big] = 0.$$

Hence

$$\begin{aligned} E\Big(\int_S^T |f_n(s,Y_s)|ds\Big) &= E\Big(|u_n(T,Y_T)-u_n(S,Y_S)|\Big) \\ &\leqslant 2 \sup_{(t,x)\in[S,T]\times\mathbb{R}^d} |u_n(t,x)| \\ &\leqslant 2 \sup_{(t,x)\in[S,T]\times\mathbb{R}^d} |u(t,x)| \\ &\leqslant \|f\|_{\mathbb{L}^{\hat q}_{\hat{\mathbf{p}}}(T)}. \end{aligned}$$

By (3.42) and letting $n\to\infty$, we get (3.41).

Now by applying (3.41) to $f=b^2$, we can get

$$\mathbb{E}\left(\int_S^T |b(t,Y_t)|^2 \mathrm{d}t\right) \leqslant C\|b^2\|_{\mathbb{L}^{q/2}_{\mathbf{p}/2}(S,T)} = C\|b\|^2_{\mathbb{L}^q_{\mathbf{p}}(S,T)}.$$

Then, Khasminskii's lemma shows that for any constant $\kappa>0$,

$$\mathbb{E}\exp\left\{\kappa\int_0^T b^2(s,Y_s)\mathrm{d}s\right\} \leqslant C(\kappa,d,\mathbf{p},q)\|b\|^2_{\mathbb{L}^q_{\mathbf{p}}(T)} < \infty. \tag{3.43}$$

As a result, we have

$$\mathbb{E}\rho_T := \mathbb{E}\exp\left\{-\int_0^T \big[b^T\sigma^{-1}\big](s,Y_s)\mathrm{d}W_s - \frac{1}{2}\int_0^T \big[b^*(\sigma\sigma^*)^{-1}b\big](s,Y_s)\mathrm{d}s\right\} = 1.$$

The existence of a weak solution X_t to SDE (3.3) follows by Girsanov's theorem. Furthermore,

$$\begin{aligned} \mathbb{E}\left(\int_0^T f(s,X_s)\mathrm{d}s\right) &= E\left(\rho_T\int_0^T f(t,Y_t)dt\right) \\ &\leqslant (E\int_0^T \rho_T^\alpha dt)^{1/\alpha}(E\int_S^T f^\beta(t,Y_t)dt)^{1/\beta}, \end{aligned}$$

where α, $\beta>1$ satisfying $1/\alpha+1/\beta=1$. Since

$$\begin{aligned} E\rho_T^\alpha = E\Big[&\Big(\exp(-2\alpha\int_0^T b^T(\sigma^T)^{-1}(s,Y_s)dW_s - 2\alpha^2\int_0^T (b^T(\sigma\sigma^T)^{-1}b)(s,Y_s)ds)\Big)^{1/2} \\ &\Big(\exp((4\alpha^2-\alpha)\int_0^T (b^T(\sigma\sigma^T)^{-1}b)(t,Y_t)^2dt)\Big)^{1/2}\Big], \end{aligned}$$

by Hölder inequality and the fact that exponential martingale is a supermartingale and (3.43), we get $E\rho_T^\alpha \leqslant C$. Then

$$\begin{aligned} E\int_0^T f(t,X_t)dt &\leqslant C(T)(E\int_0^T f^\beta(t,Y_t)dt)^{1/\beta} \\ &\leqslant C(d,\bar{\mathbf{p}},\bar q,\|b\|_{\mathbb{L}^q_{\mathbf{p}}(T)})\|f^\beta\|^{1/\beta}_{\mathbb{L}^{\bar q}_{\bar{\mathbf{p}}}(T)} \\ &= C(d,\bar{\mathbf{p}},\bar q,\|b\|_{\mathbb{L}^q_{\mathbf{p}}(T)})\|f\|_{\mathbb{L}^{\beta\bar q}_{\beta\bar{\mathbf{p}}}(T)} \\ &= C(d,\hat{\mathbf{p}},\hat q,\|b\|_{\mathbb{L}^q_{\mathbf{p}}(T)})\|f\|_{\mathbb{L}^{\hat q}_{\hat{\mathbf{p}}}(T)} \end{aligned}$$

holds by choosing β close enough to 1 such that $2/\hat{q} + 1/\hat{p}_1 + \cdots + 1/\hat{p}_d < 2/\beta$ and take $\bar{\mathbf{p}} = \hat{\mathbf{p}}/\beta$, $\bar{q} = \hat{q}/\beta$. Thus the above estimate implies (3.40). □

3.4 Itô's formula for functions in Sobolve spaces with mixed-norm

In this section, we will formulate Itô's formula for the function $u \in \mathscr{W}^q_{2,\mathbf{p}}(T)$ to the equation (3.12), $T \in [0, \infty]$.

Theorem 3.13. *Assume* **(H**a**)** *holds and* $b \in \mathbb{L}^q_{\mathbf{p}}(T)$ *with* $q, p_1, \cdots, p_d \in (2, \infty)$ *and* $2/q + 1/p_1 + \cdots + 1/p_d < 1$. $(X_t)_{t\geqslant 0}$ *is the solution to the SDE (3.3). Then there exists a version of* u *such that for* $0 \leqslant s < t \leqslant T$ *we have*

$$\begin{aligned} u(t, X_t) =& u(s, X_s) \\ &+ \int_s^t \partial_t u(r, X_r) dr + \int_s^t \nabla u(r, X_r) \cdot b(r, X_r) dr \\ &+ \int_s^t \nabla u(r, X_r) \cdot \sigma(r, X_r) dW_r \\ &+ \frac{1}{2} \sum_{i,j,l=1}^d \int_s^t \partial_{ij} u(r, X_r) \sigma_{il}(r, X_r) \sigma_{jl}(r, X_r) dr \quad a.s. \end{aligned} \tag{3.44}$$

Proof. Let $(\rho_n)_{n\in\mathbb{N}}$ be an sequence of mollifiers, define $u_n := u \star \rho_n$, then by (3.14), (3.15) and (3.16), we have u_n, $\nabla u_n \in \mathbb{L}^\infty(T)$, further we have

$$\|\partial_t u_n - \partial_t u\|_{\mathbb{L}^q_{\mathbf{p}}(T)}, \quad \|u_n - u\|_{W^q_{2,\mathbf{p}}(T)} \to 0 \text{ as } n \to \infty,$$

and

$$\|u_n - u\|_{\mathbb{L}^\infty(T)} \to 0 \text{ as } n \to \infty.$$

By classical Itô's formula for $\mathcal{C}^2$ function, we have for each $n \in \mathbb{N}$,

$$\begin{aligned} u_n(t, X_t) =& u_n(s, X_s) \\ &+ \int_s^t \partial_t u_n(r, X_r) dr + \int_s^t \nabla u_n(r, X_r) \cdot b(r, X_r) dr \\ &+ \int_s^t \nabla u_n(r, X_r) \cdot \sigma(r, X_r) dW_r \\ &+ \frac{1}{2} \sum_{i,j,l=1}^d \int_s^t \partial_{ij} u_n(r, X_r) \sigma_{il}(r, X_r) \sigma_{jl}(r, X_r) dr, \quad a.s.. \end{aligned}$$

Firstly we have

$$E|u_n(t, X_t) - u(t, X_t)| \leqslant C\|u_n - u\|_{\mathbb{L}^\infty(T)} \to 0 \text{ as } n \to \infty, \tag{3.45}$$

it also yields

$$E|u_n(s, X_s) - u(s, X_s)| \to 0 \text{ as } n \to \infty. \tag{3.46}$$

Krylov's estimate (3.40) yields

$$E\left|\int_s^t (\partial_t u_n - \partial_t u)(r, X_r)dr\right| \leqslant C\|\partial_t u_n - \partial_t u\|_{\mathbb{L}_{\mathbf{p}}^q(T)} \to 0 \text{ as } n \to \infty. \tag{3.47}$$

And

$$E\left|\int_s^t (\nabla u_n - \nabla u)(r, X_r) \cdot b(r, X_r)dr\right| \leqslant C \sup_{(t,x)\in[0,\infty)\times\mathbb{R}^d} |\nabla u_n(t,x) - \nabla u(t,x)| \|b\|_{\mathbb{L}_{\mathbf{p}}^q(T)} \to 0 \text{ as } n \to \infty. \tag{3.48}$$

Since σ is bounded, we get

$$\begin{aligned} E\left|\frac{1}{2}\sum_{i,j,l=1}^d \int_s^t (\partial_{ij}u_n - \partial_{ij}u)(r, X_r)\sigma_{il}(r, X_r)\sigma_{jl}(r, X_r)dr\right| \\ \leqslant CE\int_s^t (\partial_{ij}u_n - \partial_{ij}u)(r, X_r)dr \\ \leqslant C\|\partial_{ij}u_n - \partial_{ij}u\|_{\mathbb{L}_{\mathbf{p}}^q(T)} \to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.49}$$

By the martingale property, we have

$$E\left|\int_s^t \nabla(u_n - u)(r, X_r) \cdot \sigma(r, X_r)dW_r\right| = 0. \tag{3.50}$$

Then (3.45), (3.46), (3.47), (3.48), (3.49) and (3.50) imply

$$\begin{aligned} Eu(t, X_t) =& Eu(s, X_s) \\ &+ E\int_s^t \partial_t u(r, X_r)dr + E\int_s^t \nabla u(r, X_r) \cdot b(r, X_r)dr \\ &+ E\int_s^t \nabla u(r, X_r) \cdot \sigma(r, X_r)dW_r \\ &+ \frac{1}{2}\sum_{i,j,l=1}^d E\int_s^t \partial_{ij}u(r, X_r)\sigma_{il}(r, X_r)\sigma_{jl}(r, X_r)dr. \end{aligned} \tag{3.51}$$

Therefore there exits a subsequence $\{n_k\}_{k\in\mathbb{N}}$ such that

$$u_{n_k}(t, X_t) \to u(t, X_t) \quad a.s.$$

and

$$
\begin{aligned}
& u_{n_k}(s, X_s) \\
& + \int_s^t \partial_t u_{n_k}(r, X_r) dr + \int_s^t \nabla u_{n_k}(r, X_r) \cdot b(r, X_r) dr \\
& + \int_s^t \nabla u_{n_k}(r, X_r) \cdot \sigma(r, X_r) dW_r \\
& + \frac{1}{2} \sum_{i,j,l=1}^d \int_s^t \partial_{ij} u_{n_k}(r, X_r) \sigma_{il}(r, X_r) \sigma_{jl}(r, X_r) dr \\
& \overset{n\to\infty}{\longrightarrow} \\
& u(s, X_s) \\
& + \int_s^t \partial_t u(r, X_r) dr + \int_s^t \nabla u(r, X_r) \cdot b(r, X_r) dr \\
& + \int_s^t \nabla u(r, X_r) \cdot \sigma(r, X_r) dW_r \\
& + \frac{1}{2} \sum_{i,j,l=1}^d \int_s^t \partial_{ij} u(r, X_r) \sigma_{il}(r, X_r) \sigma_{jl}(r, X_r) dr \quad a.s.,
\end{aligned}
$$

which implies that (3.44) holds for $u \in \mathscr{W}^q_{2,\mathbf{p}}(T)$.

□

3.5 Pathwise uniqueness of strong solutions

Recall that the Hardy-Littlewood maximal operator $\mathcal{M}$ is defined by

$$
\mathcal{M} f(x) := \sup_{\mathbf{r} \in (0,\infty)^d} \frac{1}{|\mathbf{B_r}|} \int_{\mathbf{B_r}} f(x+y) dy, \quad f \in L^1_{loc}(\mathbb{R}^d),
$$

where for $\mathbf{r} = (r_1, r_2, \cdots, r_d)$, $\mathbf{B_r} := \{x \in \mathbb{R}^d : |x_1| < r_1, |x_2| < r_2, \cdots, |x_d| < r_d\}$. For every $f \in C_0^\infty(\mathbb{R}^d)$, it is known that there exists a constant $C_d > 0$ such that for all $x, y \in \mathbb{R}^d$ (see [67, Lemma 2.1]),

$$
|f(x) - f(y)| \leqslant C_d |x-y| \big(\mathcal{M} |\nabla f|(x) + \mathcal{M} |\nabla f|(y) \big), \tag{3.52}
$$

and the following $L^{\mathbf{p}}(\mathbb{R}^d)$-boundness for $\mathbf{p} \in (1,\infty)^d$ holds (see [32, Theorem 4.1]):

$$
\|\mathcal{M} f\|_{L^{\mathbf{p}}} \leqslant C_d \|f\|_{L^{\mathbf{p}}}. \tag{3.53}
$$

Now, we are in the position to give the uniqueness of the solution, i.e. the proof of our main theorem:

Proof of Theorem 3.1. We only need to show the pathwise uniqueness of solutions to SDE (3.3). To this end, we first assume that $(\mathbf{H}a)$ holds, and for $q \in (2,\infty]$ and $\mathbf{p} \in (2,\infty]^d$,

$$|b|, |\nabla\sigma| \in L^q([0,T]; L^{\mathbf{p}}(\mathbb{R}^d)) \quad \text{with} \quad \frac{2}{q} + \frac{1}{p_1} + \cdots + \frac{1}{p_d} < 1.$$

By Theorem 3.4, there exists a function $u \in \mathbb{H}^q_{2,\mathbf{p}}$ solves

$$\partial_t u(t,x) + \mathscr{L}^a_2 u(t,x) + \mathscr{L}^b_1 u(t,x) + b(t,x) = 0, \quad u(T,x) = 0.$$

Define $\Phi(t,x) := x + u(t,x)$. In view of (3.16), we can choose T small such that

$$1/2 < \|\nabla\Phi^{-1}\|_{\mathbb{L}^\infty(T)} \leqslant 2. \tag{3.54}$$

Assume that SDE (3.3) admits two solutions X^1_t and X^2_t. By the Krylov's estimate (3.40), we can use Itô's formula to get that the process $Y^i_t := \Phi(t, X^i_t)$ satisfies

$$\mathrm{d}Y^i_t = \sigma(t,X^i_t)\nabla\Phi(t,X^i_t)\mathrm{d}W_t =: \Psi(t,X^i_t)dW_t, \quad i = 1,2.$$

Let $Z_t := X^1_t - X^2_t$, we have by (3.54) that

$$\mathbb{E}|Z_t|^2 \leqslant 2\mathbb{E}|Y^1_t - Y^2_t|^2 \leqslant 2\mathbb{E}\left(\int_0^t |Z_s|^2 \mathrm{d}A_s\right),$$

where

$$A_t := \int_0^t \frac{|\Psi(s,X^1_s) - \Psi(s,X^2_s)|^2}{|Z_s|^2}\mathrm{d}s.$$

Let ρ_n be a family of mollifiers on $\mathbb{R}^d$, define $\Psi^n(t,x) := \Psi(s,\cdot) * \rho^n(x)$. Then we can write

$$\begin{aligned}
\mathbb{E}A_t &\leqslant \varliminf_{\epsilon\downarrow 0} \mathbb{E}\left(\int_0^t \frac{|\Psi(s,X^1_s) - \Psi(s,X^2_s)|^2}{|Z_s|^2} \cdot 1_{\{|Z_s|>\epsilon\}}\mathrm{d}s\right) \\
&\leqslant 3\Big(\varliminf_{\epsilon\downarrow 0} \lim_{n\to\infty} \mathbb{E}\left(\int_0^t \frac{|\Psi^n(s,X^1_s) - \Psi(s,X^1_s)|^2}{|Z_s|^2} \cdot 1_{\{|Z_s|>\epsilon\}}\mathrm{d}s\right) \\
&\quad + \varliminf_{\epsilon\downarrow 0} \lim_{n\to\infty} \mathbb{E}\left(\int_0^t \frac{|\Psi^n(s,X^2_s) - \Psi(s,X^2_s)|^2}{|Z_s|^2} \cdot 1_{\{|Z_s|>\epsilon\}}\mathrm{d}s\right) \\
&\quad + \varliminf_{\epsilon\downarrow 0} \sup_{n\in\mathbb{N}} \mathbb{E}\left(\int_0^t \frac{|\Psi^n(s,X^1_s) - \Psi^n(s,X^2_s)|^2}{|Z_s|^2} \cdot 1_{\{|Z_s|>\epsilon\}}\mathrm{d}s\right) \\
&=: 3\big(\mathcal{I}_1(t) + \mathcal{I}_2(t) + \mathcal{I}_3(t)\big).
\end{aligned}$$

By the property of mollification, it is easy to see that

$$\mathcal{I}_1(t) + \mathcal{I}_2(t) \leqslant \varliminf_{\epsilon\downarrow 0} \epsilon^{-2} \lim_{n\to\infty} C\|\Psi^n - \Psi\|^2_{\mathbb{L}^\infty(T)} = 0.$$

As for the third term, we can use (3.52), the Krylov's estimate (3.40) and (3.53) to get that

$$\begin{aligned}\mathcal{I}_3(t) &\leqslant C \sup_{n\in\mathbb{N}} \mathbb{E}\left(\int_0^t \left[\mathcal{M}|\nabla\Psi^n|(s,X_s^1) + \mathcal{M}|\nabla\Psi^n|(s,X_s^2)|\right]^2 \mathrm{d}s\right)\\ &\leqslant C \sup_{n\in\mathbb{N}} \left\|\mathcal{M}|\nabla\Psi^n|\right\|^2_{\mathbb{L}^q_{\mathbf{p}}(T)} \leqslant C\|\nabla\Psi\|^2_{\mathbb{L}^q_{\mathbf{p}}(T)} < \infty.\end{aligned}$$

Hence, as a result of the stochastic Gronwall's inequality [69, Lemma 3.7], we can get $\mathbb{E}|Z_t|^2 = 0$. The pathwise uniqueness is obtained.

Now we are going to prove (3.11), i.e. the strong Feller property. The argument follows from [74, Theorem 1.1]. It is sufficient to consider the strong Feller property of the process $(Y_t)_{t\geqslant 0}$, since $X_t := \Phi^{-1}(t, Y_t)$, (3.54). We also have that

$$dY_t = \Psi(t,Y_t)dW_t, \quad Y_0 = \Phi(x) = x + u(x),$$

where $\Psi(t,Y_t) = (\nabla\Phi\cdot\sigma)(\Phi^{-1}(t,Y_t))$. By the fact that σ is bounded and estimate (3.16), we get the boundedness of Ψ as well. We give our proof into the following steps.

Step 1. To prove
For any $T > 0$, $\gamma \in \mathbb{R}$ and all $x \neq y \in \mathbb{R}^d$, we have

$$\sup_{t\in[0,T]} E\Big(|Y_t(x) - Y_t(y)|^{2\gamma}\Big) \leqslant C|x-y|^{2\gamma}$$

where $C = C(K, \delta, \mathbf{p}, q, d, \gamma, T)$.
For $x \neq y$ and $\epsilon \in (0, |x-y|)$, define

$$\tau_\epsilon := \inf\{t \geqslant 0 : |Y_t(x) - Y_t(y)| \leqslant \epsilon\}.$$

Set $Z_t^\epsilon := Y_{t\wedge\tau_\epsilon}(x) - Y_{t\wedge\tau_\epsilon}(y)$. For any $\gamma \in \mathbb{R}$, by Itô's formula, we have

$$\begin{aligned}|Z_t^\epsilon|^{2\gamma} =& |x-y|^{2\gamma} + 2\gamma\int_0^{t\wedge\tau_\epsilon} |Z_s^\epsilon|^{2(\gamma-1)}\langle Z_s^\epsilon, [\Psi(s,Y_s(x)) - \Psi(s,Y_s(y))]dW_s\rangle\\ &+ 2\gamma\int_0^{t\wedge\tau_\epsilon} |Z_s^\epsilon|^{2(\gamma-1)}\|\Psi(s,Y_s(x)) - \Psi(s,Y_s(y))\|^2 ds\\ &+ 2\gamma(\gamma-1)\int_0^{t\wedge\tau_\epsilon} |Z_s^\epsilon|^{2(\gamma-2)}|[\Psi(s,Y_s(x)) - \Psi(s,Y_s(y))]^* Z_s^\epsilon|^2 ds\\ &=: |x-y|^{2\gamma} + \int_0^{t\wedge\tau_\epsilon} |Z_s^\epsilon|^{2\gamma}\Big(\alpha(s)dW_s + \beta(s)ds\Big),\end{aligned}$$

where

$$\alpha(s) := \frac{2\gamma[\Psi(s,Y_s(x)) - \Psi(s,Y_s(y))]^* Z_s^\epsilon}{|Z_s^\epsilon|^2}$$

and

$$\beta(s) := \frac{2\gamma\|\Psi(s,Y_s(x)) - \Psi(s,Y_s(y))\|}{|Z_s^\epsilon|^2} + \frac{2\gamma(\gamma-1)|[\Psi(s,Y_s(x)) - \Psi(s,Y_s(y))]^* Z_s^\epsilon|^2}{|Z_s^\epsilon|^4}.$$

By the Doléans-Dade's exponential (cf.[54]), we have

$$|Z_s^\epsilon|^{2\gamma} = |x-y|^{2\gamma} \exp\{\int_0^{t\wedge\tau_\epsilon} \alpha(s)dW_s - \frac{1}{2}\int_0^{t\wedge\tau_\epsilon} |\alpha(s)|^2 ds + \int_0^{t\wedge\tau_\epsilon} \beta(s)ds\}.$$

Fix $T > 0$ below. Using (3.40) and as in the proof of pathwise uniqueness, we have for any $0 \leqslant s < t \leqslant T$,

$$E\Big(\int_s^t |\beta(r\wedge\tau_\epsilon)dr|\mathcal{F}_s\Big) \leqslant C\|\nabla\Psi\|^2_{\mathbb{L}^q_{\mathbf{p}}}(s,t),$$

where $C = C(K, \delta, \mathbf{p}, q, d, \gamma, T)$. Thus, by Lemma A.1, we get for any $\lambda > 0$,

$$E\exp\Big(\lambda\int_0^{t\wedge\tau_\epsilon} |\beta(s)|ds\Big) \leqslant E\exp\Big(\lambda\int_0^T |\beta(s\wedge\tau_\epsilon)ds|\Big) < +\infty.$$

Similarly we have

$$E\Big(\exp\Big(\lambda\int_0^{t\wedge\tau_\epsilon} |\alpha(s)|^2 ds\Big) < \infty, \quad \forall\lambda > 0.$$

By Novikov's criterion,

$$t \to \exp\{2\int_0^{t\wedge\tau_\epsilon} \alpha(s)dW_s - 2\int_0^{t\wedge\tau_\epsilon} |\alpha(s)|^2 ds\} =: M_t^\epsilon$$

is a continuus expential martingale. Hence by Hölder inequality, we have

$$E|Z_t^\epsilon|^{2\gamma} \leqslant |x-y|^{2\gamma}(EM_t^\epsilon)^{\frac{1}{2}}\Big(E\exp\{\int_0^{t\wedge\tau_\epsilon} |\alpha(s)|^2 ds + 2\int_0^{t\wedge\tau_\epsilon} \beta(s)ds\}\Big)^{\frac{1}{2}} \leqslant C|x-y|^{2\gamma},$$

where C is independent of ϵ and x and y. Noticing that

$$\lim_{\epsilon\downarrow 0} \tau_\epsilon = \tau := \inf\{t \geqslant 0 : Y_t(x) = Y_t(y)\},$$

by Fatou's lemma, we obtain

$$E|Y_{t\wedge\tau}(x) - Y_{t\wedge\tau}(y)|^{2\gamma} = \lim_{\epsilon\downarrow 0} |Z_t^\epsilon|^{2\gamma} \leqslant C|x-y|^{2\gamma}.$$

Letting $\gamma = -1$ yields that

$$\tau \geqslant t, \quad a.s.$$

then we get the desired result in the begining of this step.

Step 2. To prove
For any $T > 0$, $\gamma \in \mathbb{R}$ and all $x \in \mathbb{R}^d$, we have

$$E\Big(\sup_{t\in[0,T]} (1+|Y_t(x)|^2)^\gamma\Big) \leqslant C_1(1+|x|^2)^\gamma, \tag{3.55}$$

where $C_1 = C_1(K, \gamma, T)$ *and for any* $\gamma \leqslant 1$ *and* $t, s \geqslant 0$,

$$\sup_{x \in \mathbb{R}^d} E|Y_s(x) - Y_t(x)|^{2\gamma} \leqslant C_2|t-s|^\gamma \tag{3.56}$$

where $C_2 = C_2(K, \gamma)$.
Since σ is bounded, we can obtain (3.55) by Itô's formula. For (3.56), by Burkholder-Davis-Gundy inequality and the fact that σ is bounded, we have for $\gamma \in [1, \infty)$,

$$\sup_{x \in \mathbb{R}^d} E|Y_s(x) - Y_t(x)|^{2\gamma} \leqslant \sup_{x \in \mathbb{R}^d} E|\int_s^t \sigma(u, Y_u(x)) dW_s|^{2\gamma} \leqslant C_2|t-s|^\gamma.$$

Step 3. To prove
For all $t > 0$,

$$y \to Y_t(y)$$

is a homemorphism on $\mathbb{R}^d$.
For $x \neq y \in \mathbb{R}^d$, define

$$\mathcal{H}_t(x, y) =: |Y_t(x) - Y_t(y)|^{-1}.$$

For any $x, y, x', y' \in \mathbb{R}^d$ with $x \neq x'$, and $y \neq y'$, $s \neq t$, it is easy to see that

$$|\mathcal{H}_t(x, y) - \mathcal{H}_s(x', y')| \leqslant \mathcal{H}_t(x, y) \cdot \mathcal{H}_s(x', y')[|Y_t(x) - Y_s(x')| + |Y_t(y) - Y_s(y')|].$$

By *Step 1* and *Step 2*, for any $\gamma \geqslant 1$ and $s, t \in [0, T]$, we have

$$E|\mathcal{H}_t(x, y) - \mathcal{H}_s(x', y')|^\gamma \leqslant C|x-y|^{-\gamma}|x'-y'|^{-\gamma}(|t-s|^{\gamma/2} + |x-x'|^\gamma + |x-y'|^\gamma).$$

Choosing $\gamma > 4(d+1)$, by Kolmogrov's continuity criterion, there exists a continuous version to the mapping $(t, x, y) \to \mathcal{H}_t(x, y)$ on $\{(t, x, y) \in [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d : x \neq y\}$. In particular, this proves that for almost all ω, the mapping $x \to Y_t(\omega, x)$ is one-to-one for all $t \geqslant 0$.
As for the onto property, let us define

$$\phi_t(x, y) = \begin{cases} (1 + |Y_t(x|x|^{-2})|)^{-1}, & x = 0, \\ 0, & x \neq 0. \end{cases}$$

As above, using the results from *Step 2* and *Step 1*, one can show that $(t, x) \to \phi_t(x)$ admits a continuous version. Thus, $(t, x) \to Y_t(\omega, x)$ can be extended to a continuous map from $[0, \infty) \times \mathbb{R}^d \cup \{\infty\}$ to $\mathbb{R}^d \cup \{\infty\}$, where $\mathbb{R}^d \cup \{\infty\}$ is the one point compactification of $\mathbb{R}^d$. Hence $Y_t(\omega, \cdot) : \mathbb{R}^d \cup \{\infty\} \to \mathbb{R}^d \cup \{\infty\}$ is homotopic to the identity mapping $Y_0(\omega, \cdot)$ so that it is an onto map by the well known fact in homotopic theory. In particular, for almost ω, $x \to Y_t(\omega, x)$ is a homeomorphism on $\mathbb{R}^d \cup \{\infty\}$ for all $t \geqslant 0$. Clearly, the restriction of $\mathbb{R}^d$ is still a homeomorphism since $Y_t(\omega, \infty) = \infty$.
Step 4. To prove
For any bounded measurable function ϕ, $T > 0$ *and* $x, y \in \mathbb{R}^d$,

$$|E(\phi(Y_t(x))) - E(\phi(Y_t(y))) \leqslant \frac{C_T}{\sqrt{t}} \|\phi\|_\infty |x-y|, \quad \forall t \in (0, T].$$

We define $\Psi_n(t,x) := \Psi(t,\cdot) \star \rho_n(x)$, where ρ_n is a mollifier in $\mathbb{R}^d$, then we have for all $(t,x) \in [0,\infty) \times \mathbb{R}^d$,

$$\delta|\lambda|^2 \leqslant |\Psi_n^*(t,x)\lambda|^2 \leqslant K|\lambda|^2, \quad \forall \lambda \in \mathbb{R}^d.$$

Let $Y_t^n(x)$ be the unique strong solution to SDE $dY_t^n = \Psi_n(t, Y_t^n)dW_t$, $Y_0^n = Y_0$. By monotone class theorem, it suffices to prove (3.11) for any bounded Lipschitz continuous function ϕ. First of all, by Bismut-Elworthy-Li's formula (cf.[19]), for any $h \in \mathbb{R}^d$, we have

$$\nabla_h E\phi(Y_t^n(x)) = \frac{1}{t} E\Big[\phi(Y_t^n(x)) \int_0^t [\Psi_n(s, Y_s^n(x))]^{-1} \nabla_h Y_s^n(x) dW_s\Big], \tag{3.57}$$

where for a smooth function f, we denote $\nabla_h f := \langle \nabla f, h\rangle$. Noting that

$$\nabla_h Y_t^n(x) = h + \int_0^t \nabla\Psi_n(s, Y_s^n(x)) \cdot \nabla_h Y_s^n(x) dW_s,$$

by Itô's formula, we have

$$\begin{aligned}
|\nabla_h Y_t^n(x)|^2 =& |h|^2 + 2\int_0^t \langle \nabla_h Y_t^n(x), \nabla\Psi_n(s, Y_s^n(x)) \cdot \nabla_h Y_s^n(x) dW_s\rangle \\
&+ \int_0^t \|\nabla\Psi_n(s, Y_s^n(x)) \cdot \nabla_h Y_s^n(x)\|^2 ds \\
=:& \, |h|^2 + \int_0^t |\nabla_h Y_t^n(x)|^2 \Big(\alpha_h^n(s) dW_s + \beta_h^n(s) ds\Big),
\end{aligned}$$

where

$$\alpha_h^n(s) := \frac{(\nabla_h Y_t^n(x))^* \nabla\Psi_n(s, Y_s^n(x)) \cdot \nabla_h Y_s^n(x)}{|\nabla_h Y_t^n(x)|^2}$$

and

$$\beta_h^n(s) := \frac{\|\nabla\Psi_n(s, Y_s^n(x)) \cdot \nabla_h Y_s^n(x)\|^2}{|\nabla_h Y_t^n(x)|^2}.$$

By the Doléans-Dade's exponential again, we have

$$|\nabla_h Y_t^n(x)|^2 = |h|^2 \exp\{\int_0^t \alpha_h^n(s) dW_s - \frac{1}{2}\int_0^t |\alpha_h^n(s)|^2 ds + \int_0^t \beta_h^n(s) ds.\}$$

Fix $T > 0$. By (3.40), we have for any $0 \leqslant s < t \leqslant T$,

$$E\Big(\int_s^t |\beta_h^n(s)| \mathcal{F}_s\Big) \leqslant C\|\nabla\Psi_n\|^2_{\mathbb{L}^q_{\mathbf{p}}(s,t)} \leqslant C\|\nabla\Psi\|^2_{\mathbb{L}^q_{\mathbf{p}}(s,t)},$$

where $C = C(K, \delta, p, q, d, T)$ is independent of n, x and h. Thus by Lemma A.1 we get for any $\lambda > 0$,

$$\sup_n \sup_{h \in \mathbb{R}^d} E\exp\Big(\lambda \int_0^T |\beta_h^n(s)| ds\Big) < \infty.$$

With the same argument we get

$$\sup_n \sup_{h \in \mathbb{R}^d} E \exp\Big(\lambda \int_0^T |\alpha_h^n(s)| ds\Big) < \infty.$$

Hence,

$$\sup_n \sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} E|\nabla_h Y_s^n(x)|^2 \leqslant C|h|^2, \quad \forall h \in \mathbb{R}^d,$$

and by boundedness of Ψ and (3.57),

$$\begin{aligned} |\nabla_h E\phi(Y_s^n(x))| &\leqslant \frac{\|\phi\|_\infty}{t}\Big(E\int_0^t |[\Psi_n(s, Y_s^n(x))]^{-1}\nabla_h Y_s^n(x)|^2 ds\Big)^{\frac{1}{2}} \\ &\leqslant \frac{C_T\|\phi\|_\infty}{t}\Big(E\int_0^t |\nabla_h Y_s^n(x)|^2 ds\Big)^{\frac{1}{2}} \\ &\leqslant \frac{C_T\|\phi\|_\infty |h|}{\sqrt{t}}, \end{aligned}$$

which implies that for all $t \in [0, T]$ and $x, y \in \mathbb{R}^d$,

$$|E(\phi(Y_t^n(x))) - E(\phi(Y_t^n(y))) \leqslant \frac{C_T}{\sqrt{t}}\|\phi\|_\infty |x - y|, \tag{3.58}$$

where C_T is independent of n.
Now for the completeness of our proof, we only need to take limit for (3.58) by proving that for any $x \in \mathbb{R}^d$,

$$\lim_{n\to\infty} E|Y_t^n(x) - Y_t(x)| = 0. \tag{3.59}$$

Set

$$Z_t^n(x) := Y_t^n(x) - Y_t(x)$$

and

$$\eta_n(s) := \Big(\mathcal{M}|\nabla\Psi_n|(s, Y_s^n(x)) + \mathcal{M}|\nabla\Psi_n|(s, Y_s(x))\Big).$$

For any $\lambda > 0$, by Itô's formula, we have

$$
\begin{aligned}
E|Z_t^n(x)|^2 \exp(-\lambda \int_0^t \eta_n(s)ds)& \\
=&E \int_0^t \|\Psi_n(s, Y_s^n(x)) - \Psi(s, Y_s(x))\|^2 \exp(-\lambda \int_0^s \eta_n(r)dr) \\
&- \lambda E \int_0^t \eta_n(s)|Z_s^n(x)|^2 \exp(-\lambda \int_0^s \eta_n(r)dr) \\
\leqslant& E \int_0^t \|\Psi_n(s, Y_s^n(x)) - \Psi_n(s, Y_s(x))\|^2 \exp(-\lambda \int_0^s \eta_n(r)dr) \\
&+ E \int_0^t \|\Psi_n(s, Y_s(x)) - \Psi(s, Y_s(x))\|^2 \exp(-\lambda \int_0^s \eta_n(r)dr) \\
&- \lambda E \int_0^t \eta_n(s)|Z_s^n(x)|^2 \exp(-\lambda \int_0^s \eta_n(r)dr) \\
\leqslant& (C_d - \lambda)E \int_0^t \eta_n(s)|Z_s^n(x)|^2 \exp(-\lambda \int_0^s \eta_n(r)dr) \\
&+ E \int_0^t \|\Psi_n(s, Y_s^n(x)) - \Psi_n(s, Y_s(x))\|^2 ds.
\end{aligned}
$$

Then by (3.40), we obtain that for any $\lambda \geqslant C_d$,

$$\lim_{n\to\infty} E|Z_t^n(x)|^2 \exp(-\lambda \int_0^t \eta_n(s)ds \leqslant \lim_{n\to\infty} \|\Psi_n - \Psi\|_{\mathbb{L}^q_{\mathbf{p}}(T)} = 0.$$

Furthermore, by (3.40), (3.53) and Lemma A.1, we have

$$\sup_n E \exp\left(\lambda \int_0^T \eta_n(s)ds\right) < \infty, \quad \forall \lambda, T > 0.$$

Hence by Hölder's inequality,

$$\lim_{n\to\infty} E|Z_t^n(x)| \leqslant \left[\left(E \exp\left(\lambda \int_0^T \eta_n(s)ds\right)\right)^{\frac{1}{2}} \left(E|Z_t^n(x)|^2 \exp(-\lambda \int_0^t \eta_n(s)ds)\right)^{\frac{1}{2}}\right] = 0,$$

which yields (3.11). Now we proved all of the desired results. □

4 Existence and uniqueness of a maximally defined local strong solution to an SDE driven by continuous noise in mixed-norm Lebesgue spaces on a general space time domain $Q \subset [0, \infty) \times \mathbb{R}^d$

There are several interesting situations arising from diffusions in random media and particle systems (see [37] and references therein) that the studied domain Q of equation is not the full space $\mathbb{R}_+ \times \mathbb{R}^d$ but a subdomain (e.g. $Q = \mathbb{R}^{d+1} \backslash \{x \in \mathbb{R}^{d+1} : |x| \leqslant \rho\}$, $\rho > 0$). In order to deal with this kind of situation in applications, it is important to extend the obtained result in Chapter 3 to the case where the studied domain Q is not necessary the whole space but just a subset of $\mathbb{R}_+ \times \mathbb{R}^d$.

4.1 Preliminaries and main results

Let Q be an open subset of $\mathbb{R}_+ \times \mathbb{R}^d$ and Q^n, $n \geqslant 1$, be bounded open subsets of Q such that $\overline{Q^n} \subset Q^{n+1}$ and $\cup_n Q^n = Q$. We add an object $\partial \notin Q$ to Q and define the neighborhoods of ∂ as the complements in Q of closed bounded subsets. Then $Q' = Q \cup \partial$ becomes a compact topological space, which is just the Alexandrov compactification of Q. For $\mathbf{p} = (p_1, \cdots, p_d) \in [0, \infty)^d$, $q \in [0, \infty)$ and $0 \leqslant S < T < \infty$, we denote $\mathbb{L}^q_{\mathbf{p}}(S, T)$ the space of all real Boreal measurable functions on $[S, T] \times \mathbb{R}^d$ with the norm

$$\|f\|_{\mathbb{L}^q_{\mathbf{p}}(S,T)} =: \left(\int_S^T \left(\int_{\mathbb{R}} \cdots \left(\int_{\mathbb{R}} |f(t, x_1, ..., x_d)|^{p_1} \mathrm{d}x_1 \right)^{\frac{p_2}{p_1}} \mathrm{d}x_2 \right)^{\frac{p_3}{p_2}} \cdots \mathrm{d}x_d \right)^{\frac{1}{p_d}} dt \right)^{1/q} < +\infty.$$

For simplicity, we write

$$\mathbb{L}^q_{\mathbf{p}} = \mathbb{L}^q_{\mathbf{p}}(0, \infty), \quad \mathbb{L}^q_{\mathbf{p}}(T) = \mathbb{L}^q_{\mathbf{p}}(0, T), \quad \mathbb{L}^{q,loc}_{\mathbf{p}} = L^{loc}_q(\mathbb{R}_+, L_{\mathbf{p}}(\mathbb{R}^d)).$$

The following theorem is the main result that we want to prove.

Theorem 4.1. *Assume that for any $n \in \mathbb{N}$ and some $\mathbf{p}(n) = (p_1(n), \cdots, p_d(n))$, $q(n) \in (2, \infty)$ satisfying $1/p_1(n) + \cdots + 1/p_d(n) + 2/q(n) < 1$,*
(i)$|bI_{Q^n}|$, $|\nabla \sigma I_{Q^n}| \in \mathbb{L}^{q(n)}_{\mathbf{p(n)}}$,
(ii)$\sigma_{ij}(t, x)$ is uniformly continuous in x uniformly with respect to t for $(t, x) \in Q^n$, and there exist positive constants δ_n such that for all $(t, x) \in Q^n$,

$$|\sigma(t, x)^* \lambda|^2 \geqslant \delta_n |\lambda|^2, \quad \forall \lambda \in \mathbb{R}^d.$$

Then for any $(s, x) \in Q$, there exists a unique continuous Q'−valued function $z_t =: (t, X_t)$ and a $\mathcal{F}_t$−stopping time $\xi =: inf\,\{t \geqslant 0 : z_t \notin Q\}$ such that X_t is the unique strong solution to the following SDE

$$X_t = x + \int_0^t b(s + r, X_r) dr + \int_0^t \sigma(s + r, X_r) dW_r, \quad \forall t \in [0, \xi), a.s. \tag{4.1}$$

and for any $t \geqslant 0$, $z_t = \partial$ on the set $\{\omega : t \geqslant \xi(\omega)\}$ (a.s.).

4.2 Proof of Theorem 4.1

Now we are going to prove the local well-posedness result on an arbitrary domain $Q \subset \mathbb{R}_+ \times \mathbb{R}^d$ by applying the localization technique, which is a modification of the proof of Theorem 1.3 in [73]. Furthermore we will give a precise description about the continuity of the solution on the domain Q, especially around the boundary ∂Q.

Proof. By Lemma A.4, for each $n \in \mathbb{N}$, we could find a nonnegative smooth function $\chi_n(t,x) \in [0,1]$ in $\mathbb{R}^{d+1}$ such that $\chi_n(t,x) = 1$ for all $(t,x) \in Q^n$ and $\chi_n(t,x) = 0$ for all $(t,x) \notin Q^{n+1}$. Let

$$b_s^n(t,x) =: \chi_n(t+s,x)b(t+s,x)$$

and

$$\sigma_s^n(t,x) =: \chi_{n+1}(t+s,x)\sigma(t+s,x) + (1-\chi_n(t+s,x))(1+\sup_{(t+s,x)\in Q^{n+2}}|\sigma(t+s,x)|)\mathbb{I}_{d\times d}.$$

By Theorem 3.1 there exists a unique strong continuous solution X_t^n satisfying

$$X_t^n = x + \int_0^t b^n(r,X_r^n)dr + \int_0^t \sigma^n(r,X_r^n)dW_r, \quad \forall t\in[0,\infty), a.s. \tag{4.2}$$

More precisely, for conditions in Theorem 3.1, for any $(t,x)\in\mathbb{R}^d$,

$$|b_s^n(t,x)| \leqslant |(bI_{Q^{n+2}})(s+t,x)|,$$

$$\begin{aligned}|\nabla\sigma_s^n(t,x)| &\leqslant |(\nabla\chi_{n+1}\sigma)(t+s,x)| + |(\chi_{n+1}\nabla\sigma)(t+s,x)| + c|\nabla\chi_n(t+s,x)|\\ &\leqslant |(\nabla\chi_{n+1}\sigma I_{Q^{n+2}})(t+s,x)| + |(\nabla\sigma I_{Q^{n+2}})(t+s,x)| + c|\nabla\chi_n(t+s,x)|,\end{aligned}$$

which means we can take $p =: p_{n+2}$, $q =: q_{n+2}$. For condition (ii), $\sigma(t,x)$ is uniformly continuous in x uniformly with respect to t for $(t,x)\in Q^{n+3}$, then $(\chi_{n+1}\sigma)(s+t,x)$ is uniformly continuous in $x\in\mathbb{R}^d$ locally uniformly with respect to $t\in\mathbb{R}$, and σ^n as well. Further there exist constants $K(n)$ and $\delta(n)$ such that for all $(t,x)\in\mathbb{R}^{d+1}$, and $\forall\lambda\in\mathbb{R}^d$,

$$|\sigma_s^n(t,x)\lambda|^2 \leqslant |(\sigma I_{Q^{n+2}} + (1+\sup_{(s+t,x)\in Q^{n+2}}|\sigma(s+t,x)|)I_{d\times d})(s+t,x)\lambda|^2 \leqslant K(n)|\lambda|^2,$$

and

$$\begin{aligned}|\sigma_s^n(t,x)\lambda|^2 \geqslant &|(\sigma I_{Q^{n+1}} + I_{(Q^{n+1})^c\cap Q^{n+2}} + I_{(Q^{n+2})^c}(1\\ &+\sup_{(s+t,x)\in Q^{n+2}}|\sigma(s+t,x)|)I_{d\times d})(s+t,x)\lambda|^2\\ \geqslant&(\delta(n)\wedge 1)|\lambda|^2.\end{aligned}$$

Thus conditions in Theorem 3.1 are fulfilled. For $n\geqslant k$, define

$$\tau_{n,k} =: \inf\left\{t\geqslant 0 : z_t^n =: (s+t,X_t^n)\notin Q^k\right\},$$

then it is easy to see X_t^n, X_t^k satisfy

$$X_{t\wedge\tau_{n,k}} = x + \int_0^{t\wedge\tau_{n,k}} b_s^k(r, X_r)dr + \int_0^{t\wedge\tau_{n,k}} \sigma_s^k(r, X_r)dW_r, \quad a.s.$$

By the local uniqueness of the solution in Theorem 3.1, we have

$$P\left\{\omega : X_t^n(\omega) = X_t^k(\omega), \forall t \in [0, \tau_{n,k}(\omega))\right\} = 1,$$

which implies $\tau_{k,k} \leqslant \tau_{n,k} \leqslant \tau_{n,n}$ $a.s.$. Thus if we take $\xi_k =: \tau_{k,k}$, then ξ_k is an increasing sequence of stopping times, and

$$P\left\{\omega : X_t^n(\omega) = X_t^k(\omega), \forall t \in [0, \xi_k(\omega))\right\} = 1.$$

Now for each $k \in \mathbb{N}$, the definitions

$$X_t(\omega) =: X_t^k(\omega) \text{ for } t < \xi_k, \quad \xi =: \lim_{k\to\infty} \xi_k,$$

and

$$z_t = (t, X_t), \quad t < \xi, \quad z_t = \partial, \quad \xi \leqslant t < \infty$$

make sense almost surely. We may throw the set of ω where the above definitions do not make sense and work only on the remaining part of Ω. Then X_t satisfies SDE (4.1) and ξ is the related explosion time.

The next thing is to prove that z_t is continuous on Q'. Since z_t coincides with (t, X_t^n) on any Q^n before ξ_n, the continuity on any Q^n of z_t follows from the continuity of (t, X_t^n), which can be obtained by Theorem 3.1. So we only need to show that z_t is left continuous at ξ $(a.s.)$. The argument essentially follows from [37]. we first need to prove the following lemma in order to show that $(z_t)_{t\geqslant 0}$ has the strong Markov property. In the following we use $P_{s,x}^n$ to denote the distribution of process $(z_t^n)_{t\geqslant 0} = (z_t^n(s,x))_{t\geqslant 0} := (s+t, X_t^n(0,x))_{t\geqslant 0}$ on $\mathcal{C}([0,\infty), \mathbb{R}^{d+1})$, where $(X_t^n(0,x))_{t\geqslant 0}$ means the process $(X_t^n)_{t\geqslant 0}$ defined above with initial point $(0,x) \in \mathbb{R}^{d+1}$. $E_{s,x}^n$ denotes the expectation corresponding to $P_{s,x}^n$.

The following argument is based on Proposition 4.3.3 of [46].

Define the space $\mathbb{W}_0 := \left\{w \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d) | w(0) = 0\right\}$ equipped with the supremum norm and Borel $\sigma-$algebra $\mathcal{B}(\mathbb{W}_0)$, the class $\mathcal{E}$ collects all the maps $F : \mathbb{R}^d \times \mathbb{W}_0 \to \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ such that for every probability measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ there exists a $\overline{\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{W}_0)}^{\mu\times P^W}$ $/\mathcal{B}(\mathbb{R}^d)$ measurable map $F_\mu : \mathbb{R}^d \times \mathbb{W}_0 \to \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ such that for $\mu - a.e. x \in \mathbb{R}^d$ we have $F(x,w) = F_\mu(x,w)$ for $P^W - a.e.$ $w \in \mathbb{W}_0$. Here $\overline{\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{W}_0)}^{\mu\times P^W}$ means the completion of $\mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{W}_0)$ with respect to $\mu \times P^W$, and P^W denotes the distribution of the standard d-dimensional Wiener process $(W_t)_{t\geqslant 0}$ on $(\mathbb{W}_0, \mathcal{B}(\mathbb{W}_0))$. For each $n \in \mathbb{N}$, since we already have the pathwise uniqueness and existence of strong solution $(X_t^n)_{t\geqslant 0}$ to (4.2), by applying Theorem E.8 in [46], we obtain that there exists a map $F \in \mathcal{E}$ such that for $u \leqslant t$ we have $X_t^n(s,(0,x))(\omega) = F_{P\circ(X_u^n(s,(0,x)))^{-1}}(X_u^n(s,(0,x))(\omega), (W_\cdot - W_u)(\omega))(t)$ for

$P-a.e.\ \omega \in \Omega$. Then for every bounded measurable function f and all $u, t \in [0,\infty)$ with $u \leqslant t$ we have for $P-a.e.\ \omega \in \Omega$

$$\begin{aligned} E[f(X_t^n(s,(0,x)))|\mathcal{F}_u](\omega) &= E[f(F_{P\circ(X_u^n(s,(0,x)))^{-1}}(X_u^n(s,(0,x))(\omega), W_\cdot - W_u)(t))] \\ &= E[f(F_{\delta_{X_u^n(s,(0,x))(\omega)}}(X_u^n(s,(0,x))(\omega), W_\cdot - W_u)(t))] \\ &= E[f(X_t^n(s,(u,X_u^n(s,(0,x))))(\omega))], \end{aligned} \tag{4.3}$$

which shows the Markov property of the process $(X_t^n)_{t\geqslant 0}$. Here $X_t^n(s,(u,X_u^n(s,(0,x))))$ means the solution $(X_t^n)_{t\geqslant 0}$ to (4.2) with starting point $(u, X_u^n(s,(0,x))) \in \mathbb{R}^{d+1}$. Combining with the Feller property of $(X_t^n)_{t\geqslant 0}$ yielding from the second statement of Theorem 3.1 and well known results about Markov processes (see e.g. [11, Theorem 16.21]), we get that $(X_t^n)_{t\geqslant 0}$ is a strong Markov process.

Now we are going to prove that $(z_t^n)_{t\geqslant 0}$ is a strong Markov process. Observing that for $u \geqslant 0$, $(\hat{W}_t)_{t\geqslant 0} := (W_{t+u} - W_u)_{t\geqslant 0}$ is still a Brownian motion. For any $(s,x) \in Q$, and for any Borel bounded function f on $\mathbb{R}^{d+1}$, by (4.3), we have for any $u, t \geqslant 0$, $P-a.e.$

$$\begin{aligned} &X_{t+u}^n(s,(u,X_u^n(s,(0,x)))) \\ &\quad = X_u^n(s,(0,x)) + \int_u^{u+t} \sigma_s^n(r, X_r^n(s,(0,x)))d(W_r - W_u) \\ &\qquad\qquad + \int_u^{u+t} b_s^n(r, X_r^n(s,(0,x)))dr \\ &\quad = X_u^n(s,(0,x)) + \int_0^t \sigma_s^n(r+u, X_{u+r}^n(s,(0,x)))d\hat{W}_r \\ &\qquad\qquad + \int_0^t b_s^n(r+u, X_{r+u}^n(s,(0,x)))dr, \end{aligned}$$

and

$$\begin{aligned} &X_t^n(s+u,(0,X_u^n(s,(0,x)))) \\ &\quad = X_u^n(s,(0,x)) + \int_0^t \sigma_s^n(u+r, X_r^n(u+s,(0,X_u^n(s,(0,x)))))d\hat{W}_r \\ &\qquad\qquad + \int_0^t b_s^n(u+r, X_r^n(u+s,(0,X_u^n(s,(0,x)))))dr \\ &\quad = X_u^n(s,(0,x)) + \int_0^t \sigma_{s+u}^n(r, X_r^n(u+s,(0,X_u^n(s,(0,x)))))d\hat{W}_r \\ &\qquad\qquad + \int_0^t b_{s+u}^n(r, X_r^n(u+s,(0,X_u^n(s,(0,x)))))dr. \end{aligned}$$

Since $\sigma_s^n(u+r,\cdot) = \sigma_{s+u}^n(r,\cdot)$, and $b_s^n(u+r,\cdot) = b_{s+u}^n(r,\cdot)$, by the pathwise uniqueness of the the following equation

$$dX_t = \sigma_{s+u}^n(t,X_t)d\hat{W}_t + b_{s+u}^n(t,X_t)dt, \quad X_0 = X_u^n(s,(0,x)),$$

we have for arbitrary Borel bounded function h on $\mathbb{R}^d$, $Eh(X^n_{t+u}(s,(u,X^n_u(s,(0,x))))) = Eh(X^n_t(s+u,(0,X^n_u(s,(0,x)))))$. Hence for $P-a.e.\ \omega \in \Omega$,

$$\begin{aligned} E[f(z^n_{t+u}(s,x))|\mathcal{F}_u](\omega) &= E[f(s+t+u, X^n_{t+u}(s,(0,x)))|\mathcal{F}_u](\omega) \\ &= E[f(s+t+u, X^n_{t+u}(s,(u,X^n_u(s,(0,x)))))(\omega))] \\ &= E[f(s+t+u, X^n_t(s+u,(0,X^n_u(s,(0,x)))))(\omega))] \\ &= E^n_{z^n_u(s,x)(\omega)} f(z^n_t). \end{aligned}$$

So $(z^n_t)_{t\geqslant 0}$ is a Markov process. Furthermore, for any $(s,x) \in Q$, by applying Ito's formula to process $X^n_r(s,(0,x))$, we get that $u^n_s(t,x) = Ef(X^n_t(s,(0,x)))$ is the solution to the following equation

$$\begin{cases} D_r u^n_s(r,x) = \dfrac{1}{2}\displaystyle\sum_{i,j=1}^d a^n_{s,ij}(r,x)\partial_i\partial_j u^n_s(r,x) + b^n_s(r,x)\cdot\nabla u^n_s(r,x) \text{ on } (0,\infty)\times\mathbb{R}^d, \\ u^n_s(0,x) = f(x), \end{cases} \tag{4.4}$$

with $(a^n_{s,ij})_{1\leqslant i,j\leqslant d} = \sigma^n_s \cdot (\sigma^n_s)^*$, and Borel bounded continuous function f on $\mathbb{R}^d$. Let $u^n(t,x)$ be the solution to the following equation

$$\begin{cases} D_r u^n(r,x) = \dfrac{1}{2}\displaystyle\sum_{i,j=1}^d a^n_{ij}(r,x)\partial_i\partial_j u^n(r,x) + b^n(r,x)\cdot\nabla u^n(r,x) \text{ on } (s,\infty)\times\mathbb{R}^d, \\ u^n(s,x) = f(x), \end{cases} \tag{4.5}$$

with $(a^n_{ij})_{1\leqslant i,j\leqslant d} = \sigma^n \cdot (\sigma^n)^*$, and σ^n and b^n are defined as following

$$b^n(r,x) := b^n_0(r,x), \quad \sigma^n(r,x) := \sigma^n_0(r,x).$$

Then it is easy to see that $u^n(s+t,x)$ also satisfies (4.4), which by using uniqueness of solution to (4.4) implies $u^n_s(t,x) = u^n(s+t,x) = Ef(X^n_t(s,(0,x)))$. By Remark 10.4 [37] (or see Theorem 3.1 [67]),we know that the unique solution $u^n(t,x)$ to the above equation (4.5) has a version $u^n_*(t,x)$ which is continuous on $t \in [0,\infty)$. Then for any $s \in [0,\infty)$, we have $u^n_*(s+t,x) = Eh(X^n_t(s,(0,x)))$ $a.e.$ on $(t,x) \in [0,\infty)\times\mathbb{R}^d$. Combining with the Feller property of $(X^n_t)_{t\geqslant 0}$ yielding from the second statement of Theorem 3.1, we obtain that for $a.e.\ t \in [0,\infty)$, $u^n_*(s+t,x) = Eh(X^n_t(s,(0,x)))$ holds for all $x \in \mathbb{R}^d$. Following from the fact that $Eh(X^n_t(s,(0,x)))$ is continuous with respect to $t \in [0,\infty)$, then we get that $u^n_*(s+t,x) = Eh(X^n_t(s,(0,x)))$ holds for all $(t,x) \in [0,\infty)\times\mathbb{R}^d$, which yields the continuity of $Eh(X^n_t(s,(0,x)))$ with respect to $(s,x) \in [0,\infty)\times\mathbb{R}^d$ from u^n_*. Combining dominated convergence theorem we get that for any Borel bounded continuous function g on $\mathbb{R}^{d+1}$, and for any $(s,x) \in [0,\infty)\times\mathbb{R}^d$

$$\begin{aligned} \lim_{(u,y)\to(s,x)} E^n_{u,y} g(z^n_t) &= \lim_{(u,y)\to(s,x)} Eg(u+t, X^n_t(u,(0,y))) \\ &= \lim_{(u,y)\to(s,x)} Eg(s+t, X^n_t(u,(0,y))) \\ &\quad + \lim_{(u,y)\to(s,x)} \Big(Eg(u+t, X^n_t(u,(0,y))) - Eg(s+t, X^n_t(u,(0,y)))\Big) \\ &= Eg(t+s, X^n_t(s,(0,x))). \end{aligned}$$

It shows that $(z_t^n)_{t\geqslant 0}$ also has Feller property, hence $(z_t^n)_{t\geqslant 0}$ is a strong Markov process. Then for any $(s,x)\in Q$, for any $(\mathcal{F}_t)$-adapted stopping time η and for any Borel bounded function f on $\mathbb{R}^{d+1}$,

$$E_{s,x}f(z_{\eta+t}) = f(\partial) + E_{s,x}(f(z_{\eta+t}) - f(\partial))I_{\xi>\eta+t}. \tag{4.6}$$

Since

$$\begin{aligned} E_{s,x}f(z_{\eta+t})I_{\xi>\eta+t} &= \lim_{n\to\infty} E_{s,x}f(z_{\eta+t})I_{\xi_n\geqslant\eta+t} \\ &= \lim_{n\to\infty} E^n_{s,x}f(z^n_{\eta+t})I_{\xi_n\geqslant\eta+t}I_{\xi_n\geqslant\eta} \\ &= \lim_{n\to\infty} E^n_{s,x}f(\eta+t, X^n_{\eta+t})I_{\xi_n\geqslant\eta+t}I_{\xi_n\geqslant\eta}, \end{aligned}$$

and $\{\xi_n\geqslant\eta\}\subset\mathcal{F}_\eta$, by the strong Markov property of $(z_t^n)_{t\geqslant 0}$, we get

$$\begin{aligned} \lim_{n\to\infty} E^n_{s,x}I_{\xi_n\geqslant\eta}E^n_{(\eta,X^n_\eta)}f(t,X^n_t)I_{\xi_n\geqslant\eta} &= \lim_{n\to\infty} E^n_{s,x}I_{\xi_n\geqslant\eta}E^n_{z^n_\eta}f(z^n_t)I_{\xi_n\geqslant\eta} \\ &=E_{s,x}I_{\xi>\eta}E_{(\eta,X_\eta)}f(t,X_t)I_{\xi>\eta}. \\ &=E_{s,x}I_{\xi>\eta}E_{z_\eta}f(z_t)I_{\xi>\eta} \end{aligned}$$

Then (4.6) yields

$$E_{s,x}f(z_{\eta+t}) = E_{s,x}E_{z_\eta}f(z_t). \tag{4.7}$$

We can find that (4.7) also holds if we replace (s,x) with ∂. Hence we get the strong Markov property of the process $(z_t)_{t\geqslant 0}$.

In the following we will prove another two auxiliary lemmas in order to show that our solution does not bounce back deep into the interior of Q from near ∂Q too often on any finite interval of time. The proof of the following two lemmas follow from a similar argument as [37]. By shifting the origin in $\mathbb{R}^{d+1}$, without losing generality, we assume $(s,x)=(0,0)$.

Lemma 4.2. *For arbitrary $n\geqslant 0$, define $\nu_0=0$,*

$$\mu_k = \inf\left\{t\geqslant\nu_k : (t,X_t)\notin Q^{n+1}\right\}, \quad \nu_{k+1} = \inf\left\{t\geqslant\mu_k : (t,X_t)\in\overline{Q^n}\right\}. \tag{4.8}$$

Then for any $S\in(0,\infty)$ there exists a constant N, depending only on d, p, q, S, $\|bI_{Q^{n+1}}\|_{\mathbb{L}^q_{\mathbf{p}}}$, $\sup_{(t,x)\in Q^{n+1}}|\sigma(t,x)|$, and the diameter of Q^{n+1}, such that

$$\sum_{k=0}^{\infty}(E|X_{S\wedge\mu_k} - X_{S\wedge\nu_k}|^2)^2 \leqslant N, \quad \sum_{k=0}^{\infty}(E|S\wedge\mu_k - S\wedge\nu_k|^2)^2 \leqslant S^4.$$

Proof. We have $E|X_{S\wedge\mu_k} - X_{s\wedge\nu_k}|^2 \leqslant 2I_k + 2J_k$, where

$$I_k := E|\int_{S\wedge\nu_k}^{S\wedge\mu_k}\sigma(s,X_s)dW_s|^2, \quad J_k := E|\int_{S\wedge\nu_k}^{S\wedge\mu_k} b(s,X_s)ds|^2.$$

Observe that on the set $\{S \wedge \nu_k < S \wedge \mu_k\}$ we have $S \wedge \nu_k = \nu_k$ and $(\nu_k, X_{\nu_k}) \in \overline{Q^n} \subset Q^{n+1}$. Furthermore, $(t, X_t) \in Q^{n+1}$ for $S \wedge \nu_k < t < S \wedge \mu_k$, and we have

$$E|\int_{S\wedge\nu_k}^{S\wedge\mu_k} \sigma(s,X_s)dW_s|^2 = \sum_{i,j=1}^{d} E|\int_{S\wedge\nu_k}^{S\wedge\mu_k} \sigma^2(s,X_s)ds| \leqslant Cd^2 E|S\wedge\mu_k - S\wedge\nu_k|,$$
$$I_k^2 \leqslant Cd^4 E|S\wedge\mu_k - S\wedge\nu_k|^2 := Cd^4 \bar{I}_k \leqslant Cd^4 SE|S\wedge\mu_k - S\wedge\nu_k|,$$
$$\sum_{k=0}^{\infty}(E|\int_{S\wedge\nu_k}^{S\wedge\mu_k} \sigma(s,X_s)dW_s|^2)^2 \leqslant Cd^4S^2, \quad \sum_{k=0}^{\infty}(\bar{I}_k)^2 \leqslant (\sum_{k=0}^{\infty}\bar{I}_k)^2 \leqslant S^4.$$

Moreover, by Hölder's inequality we have

$$J_k \leqslant E|S\wedge\mu_k - S\wedge\nu_k| \int_{S\wedge\nu_k}^{S\wedge\mu_k} |b(s,X_s)|^2 ds, \quad J_k^2 \leqslant \bar{I}_k \bar{J}_k,$$

where

$$\bar{J}_k := E(\int_{S\wedge\nu_k}^{S\wedge\mu_k} |b(s,X_s)|^2 ds)^2.$$

Let $\tau_n =: \inf\{t \geqslant 0 : z_t \notin Q^n\}$. By the strong Markov property of z_t on Q it follows that

$$\bar{J}_k \leqslant \sup_{(s,x)\in Q^{n+1}} E_{s,x}(\int_0^{S\wedge\tau_{n+1}} |b(s+t,X_t)|^2 dt)^2,$$

Since before τ_{n+1}, $X_t = X_t^{n+1}$, we see that the latter expression will not change if we change arbitrarily b outside of Q^{n+1} only preserving the property that new b belongs to $\mathbb{L}_p^q$. We choose to let b be zero outside of Q^{n+1} and then get the desired estimate from (3.40). The lemma is proved.

□

Lemma 4.3. *We say that on the time interval $[\nu_k, \mu_k]$ the trajectory (t, X_t) makes a run from $\bar{Q}^n$ to $(Q^{n+1})^{\mathsf{c}}$ provided that $\mu_k < \infty$. Denote by $\nu(S)$ the number of runs which (t, X_t) makes from $\overline{Q^n}$ to Q^{n+1} before time S. Then for any $\alpha \in [0, 1/2)$, $E\nu^{\alpha}(S)$ is dominated by a constant N, which depends only on α, d, $\mathbf{p}$, q, S, $\|bI_{Q^{n+1}}\|_{\mathbb{L}_{\mathbf{p}}^q}$, $\sup_{(t,x)\in Q^{n+1}} |\sigma(t,x)|$, the diameter of Q^{n+1}, and the distance between the boundaries of Q^n and Q^{n+1}.*

Proof. For any integer $k \geqslant 1$

$$kP^2(\mu_{k-1} \leqslant S) \leqslant P^2(\mu_0 \leqslant S) + ... + P^2(\mu_{k-1} \leqslant S) + ... \tag{4.9}$$

Since

$$E\left\{|X_{S\wedge\mu_k} - X_{S\wedge\nu_k}|^2 + |S\wedge\mu_k - S\wedge\nu_k|^2\right\}$$
$$\geqslant E\left\{|X_{S\wedge\mu_k} - X_{S\wedge\nu_k}|^2 + |\mu_k - \nu_k|^2\right\} I_{\mu_k \leqslant S} \geqslant dist^2(\partial Q^n, \partial Q^{n+1}) P(\mu_k \leqslant S).$$

From Lemma 4.2 we see that series in (4.9) converges and its sum is bounded by a constant with proper dependence on the data. After that it only remains to note that $P(\nu(S) \geqslant k) = P(\mu_{k-1} \leqslant S)$. □

Now we go back to prove that z_t is left continuous at ξ $(a.s.)$. We denote $\nu_k(S)$ is the number of runs of z_t from $\bar{Q}^k$ to $(Q^{k+1})^{\mathsf{c}}$ before $S \wedge \xi$. For $n > k+1$ obviously, $\nu_k(S \wedge \xi_n)$ is also the number of runs that (t, X_t^n) makes from $\overline{Q^k}$ to $(Q^{k+1})^{\mathsf{c}}$ before $S \wedge \xi_n$, which increase if we increase the time interval to S. By Lemma 4.3 $E\nu_k^{1/4}(S \wedge \xi_n)$ is bounded by a constant independent of n. By Fatou's theorem $E\nu_k^{1/4}(S \wedge \xi)$ is finite. In particular, on the set $\{\omega : \xi(\omega) < \infty\}$ (a.s.) we have $\nu_k(\xi) < \infty$. The latter also holds on the set $\{\omega : \xi(\omega) = \infty\}$ because z_t is continuous on $[0, \xi)$ and Q^k is bounded. Thus $\nu_k(\xi) < \infty$ (a.s.) for any k. Since $(\xi^n, X^n_{\xi^n}) \in \partial Q^n$ we conclude that (a.s.) there can exist only finitely n such that z_t visits $\overline{Q^k}$ after exiting from Q^n. This is the same as to say that $z_t \to \partial$ as $t \uparrow \xi$ (a.s.).

About the uniqueness, if there is another continuous $Q'-$valued solution $z'_t = (s+t, X'_t)$ to equation (4.1) with explosion time ξ', furthermore for $t < \xi'$ it is $Q-$valued. Then for any $n \geqslant 1$

$$\tau^n(X'_\cdot) = \inf\{t \geqslant 0 : (s+t, X'_t) \notin Q^n\} < \xi' \tag{4.10}$$

and

$$\bar{\xi} := \lim_{n\to\infty} \tau^n(X'_\cdot) = \xi' \quad (a.s.). \tag{4.11}$$

Precisely $\bar{\xi} \leqslant \xi'$ by (4.10). On the other hand, on the set where $\bar{\xi} < \xi'$, we have $z'_{\bar{\xi}} \in Q$ since $\bar{\xi} < \xi'$, we also have $z'_{\bar{\xi}} = \partial$ since $z'_{\bar{\xi}}$ is the limit of points getting outside of any Q^n. Observe that before $\tau^n(X'_\cdot)$, X'_t also satisfies SDE (4.2), from local strong uniqueness of equation (4.2) proved by Theorem 3.1, we get $X_t^n = X'_t$ for $t \leqslant \tau^n(X'_\cdot)$, so $\tau^n(X'_\cdot) = \tau_{n,n}$ and by (4.11) we see that

$$\xi' = \bar{\xi} = \lim_{n\to\infty} \tau^n(X'_\cdot) = \lim_{n\to\infty} \tau_{n,n} = \xi \quad (a.s.),$$

which implies that for $t \leqslant \xi = \xi'$, and z'_t coincides with z_t from our above construction in the existence part. □

4.3 Examples

Example 4.4. *Consider the equation (1.1) when $d = 1$, $b(t,x) = -x^{-1}$, $\sigma(t,x) = (1+x^2)^{-1/2019}$, $Q = \mathbb{R}_+ \times (0,\infty)$, and $Q^n = (0,n) \times \{x : 1/n < x < n\}$.*

For any $(s,x) \in Q$, for any $n \in \mathbb{N}$, if we take $q(n) = \infty$ and $p(n) \in (2,\infty)$, then $1/p(n) + 2/q(n) < 1$. We can also easily check that $\|bI_{Q^n}\|_{\mathbb{L}^{\infty}_{p(n)}} < \infty$, and $\|\nabla\sigma I_{Q^n}\|_{\mathbb{L}^{\infty}_{p(n)}} < \infty$. Furthermore, $\sigma(t,x)$ is uniformly continuous in x uniformly with respect to t for $(t,x) \in Q^n$, and there exist positive constants $\delta_n (= (1+n^2)^{-1/2019})$ such that for all $(t,x) \in Q^n$,

$$|\sigma^*(t,x)\lambda|^2 \geqslant \delta_n|\lambda|^2, \quad \forall \lambda \in \mathbb{R}^d.$$

Hence by Theorem 4.1 there exists a unique local strong solution to the following equation

$$X_t = x - \int_0^t \frac{1}{X_r} dr + \int_0^t (1+X_r^2)^{-\frac{1}{2019}} dW_r.$$

Example 4.5. *If $d = 2$,we consider SDE in $Q = \mathbb{R}_+ \times \mathbb{R}^2 \backslash \left\{x^{(1)} = 0\right\}$ with $b(t,x) = x \ln |x^{(1)}| = (x^{(1)} \ln |x^{(1)}|, x^{(2)} \ln |x^{(1)}|)$ and $\sigma(t,x) = I_2 \cdot \ln(2 + |x|^2)$ in Q, $Q^n = (0,n) \times \left\{x \in \mathbb{R}^2 : 1/n < |x^{(1)}| < n, |x^{(2)}| < n\right\}$, where $x^{(i)}$ denotes the $i-th$ exponent of the vector $x \in \mathbb{R}^d$ and I_2 is the identity matrix in $\mathbb{R}^2$. Then by Theorem 4.1 there exists a unique local strong solution to the following SDE for $(s,x) \in Q$*

$$\begin{cases} X_t^{(1)} = x^{(1)} + \int_0^t X_r^{(1)} \ln |X_r^{(1)}| dr + \int_0^t \ln(2 + |X_r|^2) dW_r^{(1)}, \\ X_t^{(2)} = x^{(2)} + \int_0^t X_r^{(2)} \ln |X_r^{(1)}| dr + \int_0^t \ln(2 + |X_r|^2) dW_r^{(2)}, \end{cases}$$

which can be rewrite as

$$X_t = x + \int_0^t X_r \ln |X_r^{(1)}| dr + \int_0^t I_2 \ln(2 + X_r^2) dW_r.$$

More precisely, for $n \in \mathbb{N}$, we can take $\mathbf{p}(n) = (p_1(n), p_1(n)) \in (2,\infty)^2$ and $q(n) = \infty$, then $\|bI_{Q^n}\|_{\mathbb{L}^\infty_{\mathbf{p}(n)}} < \infty$, and $\|\nabla \sigma I_{Q^n}\|_{\mathbb{L}^\infty_{\mathbf{p}(n)}} < \infty$. Put $0 < \delta_n < \ln(2 + 2n^2)$, then condition (ii) in Theorem 4.1 also is fulfilled.

5 Non-explosion of the solutions to SDEs driven by continuous noise in mixed-norm Lebesgue spaces

Our aim in this Section is to extend the non-explosion results in [37] to the multiplicative noise case on general space-time domains Q in mixed-norm Lebesgue spaces. Besides, we also give two applications to diffusions in random media and particle systems. Both are generalizations of the examples in [37, Section 9] with multiplicative noises.

5.1 Preliminaries and main result

Let Q be an open subset of $\mathbb{R}_+ \times \mathbb{R}^d$ and Q^n, $n \geqslant 1$, be bounded open subsets of Q such that $\overline{Q^n} \subset Q^{n+1}$ and $\cup_n Q^n = Q$. We add an object $\partial \notin Q$ to Q and define the neighborhoods of ∂ as the complements in Q of closed bounded subsets. Then $Q' = Q \cup \partial$ becomes a compact topological space, which is just the Alexandrov compactification of Q. For $\mathbf{p} = (p_1, \cdots, p_d) \in [0,\infty)^d$, $q \in [0,\infty)$ and $0 \leqslant S < T < \infty$, we denote by $\mathbb{L}^q_{\mathbf{p}}(S,T)$ the space of all real Borel measurable functions on $[S,T] \times \mathbb{R}^d$ with the norm

$$\|f\|_{\mathbb{L}^q_{\mathbf{p}}(S,T)} =: \left(\int_S^T \left(\int_{\mathbb{R}} \cdots \left(\int_{\mathbb{R}} |f(t, x_1, ..., x_d)|^{p_1} \mathrm{d}x_1 \right)^{\frac{p_2}{p_1}} \mathrm{d}x_2 \right)^{\frac{p_3}{p_2}} \cdots \mathrm{d}x_d \right)^{\frac{1}{p_d}} dt \right)^{1/q} < +\infty.$$

For simplicity, we write

$$\mathbb{L}^q_{\mathbf{p}} = \mathbb{L}^q_{\mathbf{p}}(0,\infty), \quad \mathbb{L}^q_{\mathbf{p}}(T) = \mathbb{L}^q_{\mathbf{p}}(0,T), \quad \mathbb{L}^{q,loc}_{\mathbf{p}} = L^{loc}_q(\mathbb{R}_+, L_{\mathbf{p}}(\mathbb{R}^d)).$$

Let $\mathcal{C}([0,\infty), \mathbb{R}^d)$ denote the space of all continuous $\mathbb{R}^d$-valued functions defined on $[0,\infty)$, by $\mathcal{C}([0,\infty), Q')$ we denote all continuous $Q'-$valued paths, $\mathcal{C}^n_b(\mathbb{R}^d)$ denotes the set of all bounded n times continuously differentiable functions on $\mathbb{R}^d$ with bounded derivatives of all orders. Set $(a_{ij})_{1\leqslant i,j\leqslant d} := \sigma\sigma^*$, where σ^* denotes the transpose of σ. For $f \in L^1_{loc}(\mathbb{R}^d)$ we define $\partial_j f(x) := \frac{\partial f}{\partial x_j}(x)$ and $\nabla f := (\partial_i f)_{1\leqslant i\leqslant d}$ denotes the gradient of f. Here the derivatives are meant in the sense of distributions. For a real valued function $g \in \mathcal{C}^1([0,\infty))$, $D_t g$ denotes the derivative of g with respect to t. $L(\mathbb{R}^d)$ denotes all $d \times d$ real valued matrices.

As mentioned in [37], there are several interesting situations arising from applications, say diffusions in random media and particle systems, where the domain Q of SDE is not the full space $\mathbb{R} \times \mathbb{R}^d$ but a subdomain (e.g. $Q = \mathbb{R} \times (\mathbb{R}^d \backslash \gamma^\rho)$, where $\gamma^\rho = \{x \in \mathbb{R}^d | dist(x,\gamma) \leqslant \rho\}$, $\rho > 0$, and γ is a locally finite subset of $\mathbb{R}^d$), where none of the above results mentioned can be applied to get global solutions, except for the one in [37]. Moreover, Krylov and Röckner in [37] not only proved the existence and uniqueness of a maximal local strong solution of the equation on Q, but also they obtained that if $b = -\nabla\phi$, i.e., b is minus the gradient in space of a nonnegative function ϕ and if there exist a constant $K \in [0,\infty)$ and an integrable function h on Q defined as above such that the following Lyapunov conditions hold in the distributional sense

$$2D_t\phi \leqslant K\phi, \quad 2D_t\phi + \Delta\phi \leqslant he^{\epsilon\phi}, \quad \epsilon \in [0,2), \tag{5.1}$$

the strong solution does not blow up, which means $\xi = \infty$ *a.s.*. Here $D_t\phi$ denotes the derivative of ϕ with respect to t. This result can be applied to diffusions in random environment and also finite interacting particle systems to show that if the above Lyapunov conditions hold, the process does not exit from Q or go to infinity in finite time. However, [37] is restricted to the case where the equation is driven by additive noise, that is, the diffusion term is a Brownian motion.

As far as the non-explosion result is concerned, we have to take into account that having non-constant σ instead of $\mathbb{I}_{d\times d}$ in front of the Brownian motion means that we have to consider a different geometry on $\mathbb{R}^d$, and that this effects the Lyapunov function type conditions which are to replace (5.1) and also the form of the equation. In Remark 5.7 by comparing the underlying Kolmogrov operators, we explain why the SDE (5.5) should be considered and why (5.3) states the right Lyapunov type conditions which are analog to the ones in (5.1). This leads to some substantial changes in the proof of our non-explosion result in comparison with the one in [37].

Below we will give the non-explosion result of the solution in a special form of (4.1) on domain $Q \subset \mathbb{R}_+ \times \mathbb{R}^d$ under the following assumptions.

Assumption 1. ***(i)*** *The function $\phi(t,x)$ is a nonnegative continuous function defined on Q.*
(ii) *For each n there exist $\mathbf{p} = \mathbf{p}(n), q = q(n)$ satisfying*

$$p_1(n), \cdots, p_d(n), q(n) \in (2,\infty), \text{ and } 1/p_1(n) + \cdots + 1/p_d(n) + 2/q(n) < 1, \tag{5.2}$$

such that $|bI_{Q^n}|, |\nabla\sigma I_{Q^n}| \in \mathbb{L}^q_{\mathbf{p}}$.
(iii) *For each $1 \leqslant i,j \leqslant d$, $\sigma_{ij}(t,x)$ is uniformly continuous in $x \in \mathbb{R}^d$ locally uniformly with respect to $t \in \mathbb{R}_+$, and there exists a positive constant K such that for all $(t,x) \in Q$,*

$$\frac{1}{K}|\lambda|^2 \leqslant |\sigma^*(t,x)\lambda|^2 \leqslant K|\lambda|^2, \quad \forall \lambda \in \mathbb{R}^d.$$

(iv) *For some constants $K_1 \in [0,\infty)$ and $\epsilon \in [0,2)$ in the sense of distributions on Q we have*

$$2D_t\phi \leqslant K_1\phi, \quad 2D_t\phi + \sum_{i,j=1}^d \partial_j(a_{ij}\partial_i\phi) \leqslant he^{\epsilon\phi}. \tag{5.3}$$

where h is a continuous nonnegative function on Q satisfying the following condition:
(H) For any $a > 0$ and $T \in (0,\infty)$ there is an $r = r(T,a) \in (1,\infty)$ such that

$$H(T,a,r) := H_Q(T,a,r) := \int_Q h^r(t,x)I_{(0,T)}(t)e^{-a|x|^2}dtdx < \infty.$$

(v) *For all $1 \leqslant i,j \leqslant d$, for all (t,x), $(s,y) \in [0,\infty) \times \mathbb{R}^d$,*

$$|a_{ij}(t,x) - a_{ij}(s,y)| \leqslant K(|x-y| \vee |t-s|^{1/2}), \tag{5.4}$$

and for all $n \in \mathbb{N}$, for (t,x), $(s,y) \in Q^n$, there exists $C_n \in [0,\infty)$ such that

$$|\partial_j a_{ij}(t,x) - \partial_j a_{ij}(s,y)| \leqslant C_n(|x-y| \vee |t-s|^{1/2}).$$

(vi) *The function ϕ blows up near the parabolic boundary of Q, that is for any $(s,x) \in Q$, $\tau \in (0,\infty)$, and continuous bounded $\mathbb{R}^d-$valued function x_t defined on $[0,\tau)$ and such that $(s+t,x_t) \in Q$ for all $t \in [0,\tau)$ and*

$$\liminf_{t\uparrow\tau} dist((s+t,x_t),\partial Q) = 0,$$

we have

$$\limsup_{t\uparrow\tau} \phi(s+t,x_t) = \infty.$$

Remark 5.1. *Observe that $H(T,a,r) < \infty$ if h is just a constant. Moreover, Assumption 1 (iii) shows that $\sigma(t,x)$ is uniformly bounded for $(t,x) \in Q$, invertible on Q, and the inverse $\sigma^{-1}(t,x)$ is also bounded in $(t,x) \in Q$.*

Theorem 5.2. *Let Assumption 1 be satisfied and let $(W_t)_{t\geqslant 0}$ be a d-dimensional Wiener process defined on a complete probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},P)$. Then for any $(s,x) \in Q$ there exists a continuous $\mathbb{R}^d$-valued and $(\mathcal{F}_t)$-adapted random process $(X_t)_{t\geqslant 0}$ such that almost surely for all $t \geqslant 0$, $(s+t,X_t) \in Q$,*

$$X_t = x+\int_0^t \sigma(s+r,X_r)dW_r+\int_0^t(-\sigma\sigma^*\nabla\phi)(s+r,X_r)dr+\frac{1}{2}(\sum_{j=1}^d\int_0^t \partial_j a_{ij}(s+r,X_r)dr)_{1\leqslant i\leqslant d}. \tag{5.5}$$

Furthermore, for each $T \in (0,\infty)$ and $m \geqslant 1$ there exists a constant N, depending only on K, K_1, d, $\mathbf{p}(m+1)$, $q(m+1)$, ϵ, T, $\|\nabla\phi I_{Q^{m+1}}\|_{\mathbb{L}^{q(m+1)}_{\mathbf{p}(m+1)}}$, $dist(\partial Q^m,\partial Q^{m+1})$, $\sup_{Q^{m+1}}\{\phi+h\}$, and the function H, such that for $(s,x) \in Q^m$, $t \leqslant T$ we have

$$E\sup_{t\leqslant T}\exp(\mu\phi(s+t,X_t)+\mu\nu|X_t|^2) \leqslant N,$$

where

$$\mu = (\delta/2)e^{-TK_1/(2\delta)}, \quad \delta = 1/2-\epsilon/4, \quad \nu = \mu/(12KT). \tag{5.6}$$

Remark 5.3. *Obviously, the Kolmogrov operator $\mathcal{L}$ corresponding to (5.5) is given by*

$$\mathcal{L} = div(\sigma\sigma^*\nabla) - \langle\sigma^*\nabla\phi,\sigma^*\nabla\rangle, \tag{5.7}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product in $\mathbb{R}^d$. Recalling that $div\circ\sigma$ is the adjoint of the 'geometric' gradient (i.e. taking into account the geometry given to $\mathbb{R}^d$ through σ) $\sigma^*\nabla$, we see that (5.5) is the geometrically correct analog of the SDE

$$dX_t = -\nabla\phi(X_t)dt + dW_t$$

studied in [37]. So, the Laplacian Δ in [37] is replaced by the Laplace-Beltrami operator $div(\sigma\sigma^*\nabla)(=\sum_{i,j=1}^d\partial_j(a_{ij}\partial_i))$ and the Euclidean gradient ∇ in [37] is replaced by the 'geometric' gradient $\sigma^*\nabla$. Also condition (5.3) is then the exact analog of condition (5.1) above, which was assumed in [37].

In order to show that under certain conditions our solutions will not blow up, we need certain auxiliary proofs and we will show it in the later several sections.

5.2 Probabilistic representation of solutions to parabolic partial differential equations

In this subsection, we first give an implicit representation of the solution to the following backward parabolic partial differential equation with a potential term $V(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R}$,

$$\begin{cases} D_t u(t,x) + \mathcal{L}u(t,x) + V(t,x)u(t,x) = 0, \quad 0 \leqslant t \leqslant T, \\ \qquad\qquad\qquad\qquad\qquad\qquad u(T,x) = f(x). \end{cases} \tag{5.8}$$

Here $T \in (0,\infty)$ and

$$\mathcal{L}f(t,x) := \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 f}{\partial x_i \partial x_j}(t,x) + b(t,x)\cdot\nabla f(t,x), \quad u \in \mathcal{C}_b^2(\mathbb{R}^{d+1}),$$

where $(a_{ij})_{1\leqslant i,j\leqslant d} = \sigma\sigma^*$. We first give the assumptions which make the representation formula hold.

Assumption 2. *(i) $\sigma \in \mathcal{C}([0,T]\times\mathbb{R}^d)$,*
(ii) there exist positive constants K and δ such that for all $(t,x) \in [0,T]\times\mathbb{R}^d$,

$$\delta|\lambda|^2 \leqslant |\sigma^*(t,x)\lambda|^2 \leqslant K|\lambda|^2, \quad \forall\lambda \in \mathbb{R}^d,$$

(iii) b, $V \in \mathcal{C}_b([0,T]\times\mathbb{R}^d)$,
(iv) for all (t,x), $(s,y) \in [0,T]\times\mathbb{R}^d$, there exists constants C_1, C_2 and C_3 such that

$$\begin{aligned} |a_{ij}(t,x) - a_{ij}(s,y)| &\leqslant C_1(|x-y| \vee |t-s|^{1/2}), \\ |b(t,x) - b(s,y)| &\leqslant C_2(|x-y| \vee |t-s|^{1/2}), \\ |V(t,x) - V(s,y)| &\leqslant C_3(|x-y| \vee |t-s|^{1/2}). \end{aligned}$$

(v) $f \in \mathcal{C}_c^2(\mathbb{R}^d)$.

Theorem 5.4. *If Assumption 2 holds, then there exists a unique solution $u(t,x)$ to equation (5.8) and it can be represented by the following formula*

$$u(t,x) = E\Big[f(X(T,t,x))e^{\int_t^T V(u,X(u,t,x))du}\Big], \quad (t,x) \in [0,T]\times\mathbb{R}^d, \tag{5.9}$$

where $X(T,t,x)$ is the solution to the following SDE

$$X_s = x + \int_0^t b(t+r,X_r)dr + \int_0^s \sigma(t+r,X_r)dW_r, \quad s \geqslant 0, \tag{5.10}$$

with initial point (t,x). Furthermore, for $t \in [0,T)$ we have

$$u(t,\cdot), \quad D_t u(t,\cdot), \quad \nabla u(t,\cdot), \quad \nabla^2 u(t,\cdot) \in L^1(\mathbb{R}^d). \tag{5.11}$$

Proof. On one hand by classical results from partial differential equation, we know that under our assumption there exists a unique solution $u(t,x) \in \mathcal{C}^{1,2}([0,T],\mathbb{R}^d)$ to equation (5.8) (see [40, Theorem 5.1]), which can be written in the form of a potential with kernel k (see [40, (14.2)]):

$$u(t,x) = \int_{\mathbb{R}^d} k(T,y;t,x)f(y)dy, \quad (t,x) \in [0,T] \times \mathbb{R}^d$$

satisfying

$$\lim_{t\to T} u(x,t) = \lim_{t\to T} \int_{\mathbb{R}^d} k(T,y;t,x)f(y)dy = f(x),$$

and for $s = 0,1,2$ there exists a constant C such that for $0 \leqslant t < T$ (see [40, (13.1)])

$$\partial_x^s k(T,y;t,x) \leqslant C(T-t)^{-\frac{d+s}{2}} \exp\Big(-C\frac{|y-x|^2}{T-t}\Big).$$

Then for $s = 0,1,2$, for $t \in [0,T)$ we have

$$\begin{aligned}\int_{\mathbb{R}^d} |\partial_x^s u(t,x)|dx &\leqslant \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(y)\partial_x^s k(T,y;t,x)|dydx \\ &= \int_{\mathbb{R}^d} |f(y)| \int_{\mathbb{R}^d} |\partial_x^s k(T,y;t,x)|dxdy \\ &\leqslant C(T-t)^{-\frac{s}{2}} \int_{\mathbb{R}^d} |f(y)|dy < \infty,\end{aligned}$$

which implies that for $t \in [0,T)$

$$u(t,\cdot), \quad \nabla u(t,\cdot), \quad \nabla^2 u(t,\cdot) \in L^1(\mathbb{R}^d). \tag{5.12}$$

Since b is bounded, we get $D_t u(t,\cdot) \in L^1(\mathbb{R}^d)$ easily following from the equation (5.8) and (5.12). On another hand, since b and σ are bounded and continuous, by a known result (eg. see [33, IV Theorem 2.2]) we get the existence and uniqueness of the global solution X, then by [50, Theorem 8.2.1] we get that (5.9) solves equation (5.8). Hence combining these two sides we get the desired result and also (5.11) holds. □

5.3 Some auxiliary proofs

In order to show that under certain conditions our solutions will not blow up, we need some auxiliary proofs which we collect in this subsection. We fix a $T \in (0,\infty)$, $t \in [0,T]$ define

$$Q_T := (0,T) \times \mathbb{R}^d, \quad Q^{t,r} := [0,t) \times B_r.$$

Consider the SDE (5.10) in $\mathbb{R}^d$. First we recall two results from [73], which are corresponding to Theorems 3.1 and 3.12 in mixed-norm Lebesgue space.

Lemma 5.5. *([73, Theorem 1.1]) Assume that p, $q \in (0, \infty)$ satisfying $d/p + 2/q < 1$,*
(i) $|b|$, $|\nabla\sigma| \in \mathbb{L}_p^{q,loc}$,
(ii) for all $1 \leqslant i, j \leqslant d$, $\sigma_{ij}(t, x)$ is uniformly continuous in $x \in \mathbb{R}^d$ locally uniformly with respect to $t \in \mathbb{R}_+$, and there exist positive constants K and δ such that for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$,

$$\delta|\lambda|^2 \leqslant |\sigma^*(t, x)\lambda|^2 \leqslant K|\lambda|^2, \quad \forall \lambda \in \mathbb{R}^d. \tag{5.13}$$

Then for any $\mathcal{F}_t-$stopping time τ and $x \in \mathbb{R}^d$, there exists a unique strong continuous solution X_t such that

$$P\left\{\omega : \int_0^T |b(r, X_r(\omega))|dr + \int_0^T |\sigma(r, X_r(\omega))|^2 dr < \infty, \forall T \in [0, \tau(\omega))\right\} = 1, \tag{5.14}$$

and

$$X_t = x + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r, \quad \forall t \in [0, \tau), a.s. \tag{5.15}$$

which means that if there is another continuous stochastic process Y_t also satisfying (5.14) and (5.15), then

$$P\{\omega : X_t(\omega) = Y_t(\omega), \forall t \in [0, \tau)\} = 1.$$

Moreover, for almost all ω and all $t \geqslant 0$, $x \to X_t(\omega, x)$ is a homeomorphism on $\mathbb{R}^d$ and for any $t > 0$ and bounded measurable function ψ, x, $y \in \mathbb{R}^d$,

$$|E\psi(X_t(x)) - E\psi(X_t(y))| \leqslant C_t\|\psi\|_\infty|x - y|,$$

where $C_t > 0$ satisfies $\lim_{t\to 0} C_t = +\infty$.

Krylov's estimate plays a crucial role in the well-posedness proof and also our later work.

Lemma 5.6. *([73, Theorem 2.1, Theorem 2.2]) Suppose σ satisfies the condition in Lemma 5.12 and continuous process X_t satisfies (5.14) and (5.15). Fix an $\mathcal{F}_t-$stopping time τ, $T_0 > 0$,*
(1) if b is bounded measurable, for p, $q \in (1, \infty)$ with

$$\frac{d}{p} + \frac{2}{q} < 2,$$

there exists a positive constant $N = N(K, d, p, q, T_0, \|b\|_\infty)$ such that for all $f \in \mathbb{L}_p^q(T_0)$ and $0 \leqslant S < T \leqslant T_0$,

$$E\left(\int_{S\wedge\tau}^{T\wedge\tau} f(s, X_s)ds \middle| \mathcal{F}_S\right) \leqslant N\|f\|_{\mathbb{L}_p^q(S,T)}. \tag{5.16}$$

(2) if $b \in \mathbb{L}_p^q$ *provided with*

$$\frac{d}{p} + \frac{2}{q} < 1, \tag{5.17}$$

there exists a positive constant $N = N(K, d, p, q, T_0, \|b\|_{\mathbb{L}_p^q(T_0)})$ *such that for all* $f \in \mathbb{L}_p^q(T_0)$ *and* $0 \leqslant S < T \leqslant T_0$*,*

$$E\left(\int_{S\wedge\tau}^{T\wedge\tau} f(s, X_s)ds\middle|\mathcal{F}_S\right) \leqslant N\|f\|_{\mathbb{L}_p^q(S,T)}.$$

We note that actually condition $f \in \mathbb{L}_p^q(T_0)$ with $p, q \in (1, \infty)$ and $\frac{d}{p} + \frac{2}{q} < 1$ in the above Lemma 5.6 can be improved to $f \in \mathbb{L}_{\mathbf{p}'}^{q'}(T_0)$ with $\mathbf{p}' \in (1, \infty)^d, q' \in (1, \infty)$ and $\frac{1}{p_1'} + \cdots \frac{1}{p_d'} + \frac{2}{q'} < 2$ without assuming that b is bounded in Lemma 5.7 below, which we shall prove in the following lemma. Let K_0 and T_0 be some positive constants and we give the following assumption.

Assumption 3. *(i) For all* $1 \leqslant i, j \leqslant d$, $[0, \infty) \times \mathbb{R}^d \ni (t, x) \to \sigma_{ij}(t, x) \in \mathbb{R}$ *is uniformly continuous in* x *locally uniformly with respect to* $t \in [0, \infty)$*, and there exist positive constants* K *and* δ *such that for all* $(t, x) \in [0, \infty) \times \mathbb{R}^d$

$$\delta|\lambda|^2 \leqslant |\sigma^*(t, x)\lambda|^2 \leqslant K|\lambda|^2, \quad \forall\lambda \in \mathbb{R}^d. \tag{5.18}$$

And $|\nabla\sigma| \in \mathbb{L}_{\mathbf{p}}^{q,loc}$ *with* $\mathbf{p} \in (1, \infty)^d, q \in (1, \infty)$ *satisfying* $\frac{1}{p_1} + \cdots \frac{1}{p_d} + \frac{2}{q} < 2$.
(ii) $b(t, x)$ *is Borel measurable with* $\|b\|_{\mathbb{L}_{\mathbf{p}}^q} \leqslant K_0$ *and* $b(t, x) = 0$ *for* $t > T_0$.

Lemma 5.7. *Let Assumption 3 hold. Let* $(X_t)_{t\geqslant0}$ *be a continuous* $(\mathcal{F}_t)$*-adapted process such that (5.14) and (5.15) are satisfied. Then for any Borel function* $f \in \mathbb{L}_{\mathbf{p}'}^{q'}(S, T)$ *with* $\mathbf{p}' \in (1, \infty)^d, q' \in (1, \infty)$ *and* $\frac{1}{p_1'} + \cdots \frac{1}{p_d'} + \frac{2}{q'} < 2$*, and for* $0 \leqslant S < T \leqslant T_0$*, we have*

$$E\int_S^T |f(t, X_t)|dt \leqslant N(d, p', q', K, \|b\|_{\mathbb{L}_{\mathbf{p}}^q(T_0)})\|f\|_{\mathbb{L}_{\mathbf{p}'}^{q'}(S,T)}. \tag{5.19}$$

Furthermore, for any constant $\kappa \geqslant 0$ *and* $g \in \mathbb{L}_{\mathbf{p}}^q(T_0)$*,*

$$E\exp(\kappa\int_0^{T_0} |g(t, X_t)|^2dt) < \infty. \tag{5.20}$$

Proof. By Lemma 5.5 we obtain that there exists a unique $(\mathcal{F}_t)$-adapted $\mathbb{R}^d$-valued process $(M_t)_{t\geqslant0}$ such that $M_t = x + \int_0^t \sigma(s, M_s)dW_s$, $t \geqslant 0$. By (3.40) we have for any $\mathbf{p^1} = (p_1^1, \cdots, p_d^1) \in (1, \infty)^d$, $q_1 \in (1, \infty)$ satisfying

$$\frac{1}{p_1^1} + \cdots + \frac{1}{p_d^1} + \frac{2}{q_1} < 2,$$

for $0 < S < T \leqslant T_0$, and $f \in \mathbb{L}^{q_1}_{\mathbf{p_1}}(S,T)$

$$E\left(\int_S^T |f(t,M_t)|dt\middle|\mathcal{F}_S\right) \leqslant N\|f\|_{\mathbb{L}^{q_1}_{\mathbf{p_1}}(S,T)}, \tag{5.21}$$

where N depends only on d, K, $\mathbf{p}^1$, q_1, T_0. Applying (5.21) to $f = |g|^2$ we get

$$E\left(\int_S^T |g(t,M_t)|^2dt\middle|\mathcal{F}_S\right) \leqslant N\|g^2\|_{\mathbb{L}^{q/2}_{\mathbf{p}/2}(S,T)} = N\|g\|^2_{\mathbb{L}^q_{\mathbf{p}}(S,T)}.$$

By Lemma A.1, for any $\kappa \in [0,\infty)$ we have

$$E\exp(\kappa\int_0^{T_0} |g(t,M_t)|^2dt) \leqslant N(\kappa, K, d, \mathbf{p}, q, T_0, \|g\|_{\mathbb{L}^q_{\mathbf{p}}(T_0)}),$$

then

$$E\exp(\kappa\int_0^{T_0} |g(t,M_t)|^2dt) \leqslant N(\kappa, K, d, \mathbf{p}, q, T_0, \|g\|_{\mathbb{L}^q_{\mathbf{p}}(T_0)}). \tag{5.22}$$

And also

$$E\exp(\kappa\int_0^{T_0} |b(t,M_t)|^2dt) \leqslant N(\kappa, K, K_0, d, \mathbf{p}, q, T_0). \tag{5.23}$$

The integral over $(0,T_0)$ in (5.23) can be replaced with the one over $(0,\infty)$ since $b(t,x) = 0$ for $t > T_0$. Thus for any $\kappa \in [0,\infty)$

$$E\exp(\kappa\int_0^{\infty} |b(t,M_t)|^2dt) < \infty, \tag{5.24}$$

which and (5.18) implies that for any $c \in [0,\infty)$

$$E\exp(c\int_0^{\infty} (b^*(\sigma\sigma^*)^{-1}b)(t,M_t)dt) \leqslant E\exp(\frac{c}{\delta}\int_0^{\infty} |b(t,M_t)|^2dt) < \infty. \tag{5.25}$$

For $f \in \mathbb{L}^{q'}_{\mathbf{p}'}(S,T)$ with $\mathbf{p}' \in (1,\infty)^d$, $q' \in (1,\infty)$, we can choose $\beta > 1$ sufficiently close to 1 such that

$$\frac{1}{p'_1} + \cdots + \frac{1}{p'_d} + \frac{2}{q'} < \frac{2}{\beta}.$$

By Theorem 3.12 we obtain the existence of process $(X_t)_{t\geqslant 0}$ which satisfies (5.14) and (5.15). By Lemma A.3, we have

$$\begin{aligned} E\int_S^T |f(t,X_t)|dt = E\int_S^T \rho|f(t,M_t)|dt &\leqslant (E\int_S^T \rho^\alpha dt)^{1/\alpha}(E\int_S^T |f(t,M_t)|^\beta dt)^{1/\beta} \\ &\leqslant (E\int_0^{T_0} \rho^\alpha dt)^{1/\alpha}(E\int_S^T |f(t,M_t)|^\beta dt)^{1/\beta}, \end{aligned} \tag{5.26}$$

where α, $\beta > 1$ satisfying $1/\alpha + 1/\beta = 1$, and

$$\rho := \exp(-\int_0^\infty b^*(\sigma^*)^{-1}(s, M_s)dW_s - \frac{1}{2}\int_0^\infty (b^*(\sigma\sigma^*)^{-1}b)(s, M_s)ds).$$

Since

$$\begin{aligned} E\rho^\alpha = E\Big[&\Big(\exp(-2\alpha\int_0^\infty b^*(\sigma^*)^{-1}(s, M_s)dW_s - 2\alpha^2\int_0^\infty (b^*(\sigma\sigma^*)^{-1}b)(s, M_s)ds)\Big)^{1/2} \\ &\Big(\exp((2\alpha^2 - \alpha)\int_0^\infty (b^*(\sigma\sigma^*)^{-1}b)(s, M_s)ds)\Big)^{1/2}\Big], \end{aligned} \tag{5.27}$$

by Hölder's inequality and the fact that exponential martingale is a supermartingale and (5.25), we get

$$E\rho^\alpha \leqslant N. \tag{5.28}$$

Then

$$\begin{aligned} E\int_S^T |f(t, X_t)|dt &\leqslant N(T_0)(E\int_S^T |f(t, M_t)|^\beta dt)^{1/\beta} \\ &\leqslant N(d, \mathbf{p}_1, q_1, K, \|b\|_{\mathbb{L}^q_{\mathbf{p}}(T_0)})\|f^\beta\|^{1/\beta}_{\mathbb{L}^{q_1}_{\mathbf{p}_1}(S,T)} \\ &= N(d, \mathbf{p}_1, q_1, K, \|b\|_{\mathbb{L}^q_{\mathbf{p}}(T_0)})\|f\|_{\mathbb{L}^{\beta q_1}_{\beta p_1}(S,T)} \end{aligned}$$

for $1/p^1_1 + \cdots + 1/p^1_d + 2/q_1 < 2$, where $\mathbf{p}_1 = \mathbf{p}'/\beta$, $q_1 = q'/\beta$. Thus the above estimate implies (5.19).

Furthermore, according to Lemma A.3 and (5.22),

$$\begin{aligned} E\exp(\kappa\int_0^{T_0} |g(t, X_t)|^2 dt) =& E(\rho\exp(\kappa\int_0^{T_0} |g(t, M_t)|^2 dt)) \\ \leqslant& (E\rho^2)^{1/2}(E\exp(2\kappa\int_0^{T_0} |g(t, M_t)|^2 dt))^{1/2} < \infty. \end{aligned}$$

□

Lemma 5.8. *Let $b^{(i)}(t, x)$, $i = 1, 2$ satisfy Assumption 3 and let $|b^{(1)}(t, x) - b^{(2)}(t, x)| \leqslant \bar{b}(t, x)$, where $\bar{b} \in \mathbb{L}^q_{\mathbf{p}}$. Let $(X^{(i)}_t, W^{(i)}_t)$ satisfy:*

$$X^{(i)}_t = x + \int_0^t b^{(i)}(s, X^{(i)}_s)ds + \int_0^t \sigma(s, X^{(i)}_s)dW^{(i)}_s.$$

Then for any bounded Borel functions $f^{(i)}(x)$, $i = 1, 2$ given on $\mathcal{C} := \mathcal{C}([0, \infty), \mathbb{R}^d)$ we have

$$|Ef^{(1)}(X^{(1)}_\cdot) - Ef^{(2)}(X^{(2)}_\cdot)| \leqslant N(E|f^{(1)}(M_\cdot) - f^{(2)}(M_\cdot)|^2)^{1/2} + N\sup_{\mathcal{C}}|f^{(1)}|\|\bar{b}\|_{\mathbb{L}^q_{\mathbf{p}}} \tag{5.29}$$

where $M_t = \int_0^t \sigma(s, M_s)dW_s$ and N is a constant independent of f.

Proof. According to Lemma A.3 and (5.20), we know that

$$Ef^{(2)}(X_\cdot^{(2)}) = Ef^{(2)}(X_\cdot^{(1)})\overline{\rho}_\infty,$$

where $\Delta b(t, X_t^{(1)}) := b^{(2)}(t, X_t^{(1)}) - b^{(1)}(t, X_t^{(1)})$ and

$$\overline{\rho}_t := \exp(\int_0^t \Delta b^*(\sigma^*)^{-1}(s, X_s^{(1)})dW_s^{(1)} - \frac{1}{2}\int_0^t (\Delta b^*(\sigma\sigma^*)^{-1}\Delta b)(s, X_s^{(1)})ds),$$

also $E\overline{\rho}_t = 1$ by a similar argument as the proof of Lemma 3.4. Hence the left-hand side of (5.29) is less than

$$E|f^{(1)} - f^{(2)}|(X_\cdot^{(1)})\overline{\rho}_\infty + \sup_{\mathcal{C}} |f^{(1)}|E|\overline{\rho}_\infty - 1| := I_1 + I_2 \sup_{\mathcal{C}} |f^{(1)}|.$$

Since all moments of the exponential martingale $\overline{\rho}_t$ are finite by a similar argument as the proof in Lemma 3.4, we get

$$I_1^{3/2} \leqslant NE|f^{(1)} - f^{(2)}|^{3/2}(X_\cdot^{(1)})$$

and the latter is controled by the first term on the right hand side of (5.29) by a similar argument as the proof of Lemma 3.4. To estimate I_2 we use Itô's formula to get

$$\overline{\rho}_T = 1 + \int_0^T (\Delta b^*(\sigma^*)^{-1})(s, X_s^{(1)})\overline{\rho}_s dW_s^{(1)}.$$

It follows that for any $\beta > 1$

$$I_2^2 \leqslant E|\overline{\rho}_{T_0} - 1|^2 \leqslant E\int_0^{T_0} (\Delta b^*(\sigma\sigma^*)^{-1}\Delta b)(s, X_s^{(1)})\overline{\rho}_s^2 ds \tag{5.30}$$

$$\leqslant N(\int_0^{T_0} E\overline{\rho}_s^{2\beta/(\beta-1)}ds)^{1-1/\beta}(E\int_0^{T_0} \overline{b}^{2\beta}(s, X_s^{(1)})ds)^{1/\beta}. \tag{5.31}$$

To estimate the second factor we use Lemma 5.7 with $\beta > 1$ close to 1 such that $2/q + 1/p_1 + \cdots + 1/p_d < 1/\beta$. The first factor is controlled by means of $E\overline{\rho}_{T_0}^{2\beta/(\beta-1)}$. Thus the result follows. ☐

Assumption 4. *(i) ψ is a positive function on $\mathbb{R}^{d+1}$ such that $\psi \in \mathcal{C}_b^\infty(\mathbb{R}^{d+1})$,*
(ii) $|\nabla\psi| \in \mathbb{L}_{\mathbf{p}}^{q,loc}$ with $\mathbf{p}$ and q satisfying (5.2),
(iii) σ satisfies the conditions in Lemma 5.5,
(iv) for all (t,x), $(s,y) \in [0,T] \times \mathbb{R}^d$, there exists constant K_0, $K \in [0,\infty)$ such that

$$|a_{ij}(t,x) - a_{ij}(s,y)| \leqslant K(|x-y| \vee |t-s|^{1/2}),$$
$$|\partial_j a_{ij}(t,x) - \partial_j a_{ij}(s,y)| \leqslant K_0(|x-y| \vee |t-s|^{1/2}).$$

Let W_t be a $d-$dimensional Wiener process on a given complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$, denote $(a_{ij})_{1\leqslant i,j\leqslant d} = \sigma\sigma^*$. We introduce the process $Y(t,s,x)$ satisfying

$$Y(t,s,x) = x + \int_s^t \sigma(r, Y(r,s,x))dW_r + (\frac{1}{2}\sum_{j=1}^d \int_s^t \partial_j a_{ij}(r, Y(r,s,x))dr)_{1\leqslant i\leqslant d}, \tag{5.32}$$

and process $X(t,s,x)$ satisfying

$$\begin{aligned} X(t,s,x) = x + \int_s^t \sigma(r, X(r,s,x))dW_r + (\frac{1}{2}\sum_{j=1}^d \int_s^t \partial_j a_{ij}(r, X(r,s,x))dr)_{1\leqslant i\leqslant d} \\ - \int_s^t (\sigma\sigma^*\nabla\psi)(r, X(r,s,x))dr. \end{aligned}$$

Since for $1\leqslant i\leqslant d$, $\partial_j a_{ij} = \sum_{k=1}^d \sigma_{ik}(\partial_j\sigma_{jk}) + \sum_{k=1}^d (\partial_j\sigma_{ik})\sigma_{jk}$, and $|\partial\sigma| \in \mathbb{L}_{\mathbf{p}}^{q,loc}$ from Assumption 4 (iii), we get $\sum_{j=1}^d |\partial_j a_{ij}| \in \mathbb{L}_{\mathbf{p}}^{q,loc}$. Then Lemma 5.5 can be applied here to guarantee the existence and uniqueness of global strong solutions Y_t and X_t to this two corresponding SDEs under Assumption 4.

Lemma 5.9. *Let Assumption 4 be satisfied. Take a nonnegative Borel function f on $\mathbb{R}^{d+1}$. For $t\in[0,T]$ introduce*

$$\begin{aligned} \beta_T(t,x) = \exp(- \int_t^T \nabla\psi^*\sigma(s, Y(s,t,x))dW_s - \frac{1}{2}\int_t^T |\nabla\psi^*\sigma\sigma^*\nabla\psi|(s, Y(s,t,x))ds \\ - 2\int_t^T D_t\psi(s, Y(s,t,x)ds), \end{aligned}$$

$$v_T(t,x) = E\beta_T(t,x)f(T, Y(T,t,x)), \quad c(t) = \int_{\mathbb{R}^d} e^{-2\psi(t,x)}v_T(t,x)dx.$$

Then $c(t)$ is a constant for $t\in[0,T]$.

Proof. Using a standard approximation argument it suffices to prove the result for $f\in \mathcal{C}_c^\infty(\mathbb{R}^{d+1})$. First observe that by Assumption 4 (i) and (iii), we have

$$E\exp(\frac{1}{2}\int_t^T |\nabla\psi^*\sigma\sigma^*\nabla\psi|(s, Y(s,t,x))ds) < \infty.$$

Girsanov transformation yields

$$v_T(t,x) = E\exp(-\int_t^T 2D_t\psi(s, X(s,t,x))ds)f(T, X(T,t,x)).$$

Observe that from the Assumption 4 (i), (iii) and (iv), $(\frac{1}{2}\sum_{j=1}^{d}\partial_j a_{ij})_{1\leqslant i\leqslant d}$ and $-\sigma\sigma^*\nabla\psi$ is bounded and also satisfy condition (iv) in Assumption 2. By Theorem 5.4, $v_T(t,x)$ is the solution to the following Kolmogrov equation with a potential term $-2D_t\psi$:

$$\begin{cases} D_t v_T(t,x) + \frac{1}{2}\sum_{i,j=1}^{d}\partial_j(a_{ij}\partial_i v_T(t,x)) - ((\sigma\sigma^*\nabla\psi)^*\nabla v_T)(t,x) \\ \qquad\qquad\qquad\qquad - v_T(t,x)2D_t\psi(t,x) = 0, \\ \qquad\qquad\qquad v_T(T,x) = f(T,x). \end{cases} \tag{5.33}$$

And Theorem 5.4 shows that $v_T(t,\cdot)$, $D_t v_T(t,\cdot)$, $\nabla v_T(t,\cdot)$, $\nabla^2 v_T(t,\cdot)\in L^1(\mathbb{R}^d)$. Also there exists a kernel $k(T,y;t,x)$ such that

$$v_T(t,x) = \int_{\mathbb{R}^d} k(T,y;t,x)f(T,y)dy$$

where satisfying that there exists a constant C such that for $0\leqslant t<T$ ([40, 13.1])

$$D_t k(T,y;t,x) \leqslant C(T-t)^{-\frac{d+2}{2}}\exp\Big(-C\frac{|y-x|^2}{T-t}\Big).$$

Then by mean value theorem for $h\in\mathbb{R}$ with $t+h\in(0,T)$ there exists $\theta\in(0,1)$ such that

$$\begin{aligned}\frac{|k(T,y;t+h,x)-k(T,y;t,x)|}{h} &= D_t k(T,y;t+\theta h,x)\\ &\leqslant C(T-t-\theta h)^{-\frac{d+2}{2}}\exp\Big(-C\frac{|y-x|^2}{T-t-\theta h}\Big),\end{aligned}$$

then

$$\begin{aligned}\Big|\frac{v_T(t+h,x)-v_T(t,x)}{h}\Big| &\leqslant \int_{\mathbb{R}^d}\Big|\frac{k(T,y;t+h,x)-k(T,y;t,x)}{h}\Big|f(T,y)dy\\ &\leqslant C(T-t-\theta h)^{-\frac{d+2}{2}}\int_{\mathbb{R}^d}\exp\Big(-C\frac{|y-x|^2}{T-t-\theta h}\Big)f(T,y)dy\\ &\leqslant C'(T-t)^{-\frac{d+2}{2}}\int_{\mathbb{R}^d}\exp\Big(-C\frac{|y-x|^2}{T-t}\Big)f(T,y)dy\end{aligned} \tag{5.34}$$

for small h. Denote $g(t,x)=e^{-2\psi(t,x)}v_T(t,x)$, we have for $t\in[0,T)$, $t+h\in(0,T)$,

$$\begin{aligned}&\Big|\frac{g(t+h,x)-g(t,x)}{h}\Big|\\ &\quad= \Big|\frac{e^{-\psi(t+h,x)}(v_T(t+h,x)-v_T(t,x))}{h} + \frac{v_T(t,x)(e^{-\psi(t+h,x)}-e^{-\psi(t,x)})}{h}\Big|\\ &\quad\leqslant \Big|\frac{v_T(t+h,x)-v_T(t,x)}{h}\Big| + \Big|\frac{v_T(t,x)(e^{-\psi(t+h,x)}-e^{-\psi(t,x)})}{h}\Big|\\ &\quad\leqslant C'(T-t)^{-\frac{d+2}{2}}\int_{\mathbb{R}^d}\exp\Big(-C\frac{|y-x|^2}{T-t}\Big)f(T,y)dy + C''v_T(t,x)\\ &\quad=: G_T(t,x),\end{aligned}$$

the last inequality holds because of (5.34) and mean value theorem. Since $v_T(t,\cdot) \in L^1(\mathbb{R}^d)$ and

$$\int_{\mathbb{R}^d} (T-t)^{-\frac{d+2}{2}} \int_{\mathbb{R}^d} \exp\Big(-C\frac{|y-x|^2}{T-t}\Big) f(T,y)dydx \leqslant C(T-t)^{-1} \int_{\mathbb{R}^d} f(T,y)dy < \infty,$$

it yields that $G_T(t,\cdot) \in L^1(\mathbb{R}^d)$, then by dominated convergence theorem, we have

$$\lim_{h\to 0} \frac{\int_{\mathbb{R}^d} g(t+h,x) - g(t,x)dx}{h} = \lim_{h\to 0} \int_{\mathbb{R}^d} \frac{g(t+h,x)-g(t,x)}{h} dx = \int_{\mathbb{R}^d} D_t g(t,x)dx.$$

That is to say

$$D_t \int_{\mathbb{R}^d} e^{-2\psi(t,x)} v_T(t,x)dx = \int_{\mathbb{R}^d} D_t(e^{-2\psi} v_T)(t,x)dx. \tag{5.35}$$

Besides, we can write the first equation in (5.33) in an equivalent form as

$$D_t(e^{-2\psi}v_T) + \frac{1}{2}\sum_{i,j=1}^d \partial_i(e^{-2\psi}a_{ij}\partial_j v_T) = 0. \tag{5.36}$$

Now we are going to prove

$$\int_{\mathbb{R}^d} div(F)(t,x)dx := \int_{\mathbb{R}^d} \sum_{i,j=1}^d \partial_i(e^{-2\psi}a_{ij}\partial_j v_T)(t,x)dx = 0, \quad t \in [0,T).$$

Since ψ is positive, $\partial\psi$ and a_{ij} are bounded for $1 \leqslant i,j \leqslant d$, then there exists constants C_1 and C_2 such that

$$F_i = \sum_{j=1}^d e^{-2\psi}a_{ij}\partial_j v_T \leqslant C_1 \sum_{j=1}^d |\partial_j v_T|,$$

and

$$\begin{aligned} div(F) &= \sum_{i,j=1}^d \partial_i(e^{-2\psi}a_{ij}\partial_j v_T) \\ &= \sum_{i,j=1}^d (-2\partial_i\psi e^{-2\psi}a_{ij}\partial_j v_T + \partial_i a_{ij} e^{-2\psi}\partial_j v_T + e^{-2\psi}a_{ij}\partial_i\partial_j v_T) \\ &\leqslant C_2 \sum_{i,j=1}^d (|\partial_j v_T| + |\partial_i\partial_j v_T|). \end{aligned}$$

Following from (5.11) we know that $F(t,\cdot)$ and $divF(t,\cdot)$ are L^1-integrable on $\mathbb{R}^d$ for any $t \in [0,T)$. For $n \in \mathbb{N}$, take smooth function χ_n on $\mathbb{R}^d$ such that $\chi_n(x) = 1$ when

$|x| \leqslant n$ and $\chi_n(x) = 0$ when $|x| > n+2$. Then by dominated convergence theorem and integration by parts formula for $t \in [0,T)$

$$\int_{\mathbb{R}^d} div(F)(t,x)dx = \lim_{n\to\infty}\int_{\mathbb{R}^d}\chi_n(x)div(F)(t,x)dx = -\lim_{n\to\infty}\int_{\mathbb{R}^d}\nabla\chi_n(x)\cdot F(t,x)dx = 0.$$

Hence from (5.36) and (5.35) we get

$$D_t\int_{\mathbb{R}^d} e^{-2\psi(t,x)}v_T(t,x)dx = 0.$$

This yields that $c(t)$ is a constant for $t \in [0,T)$. Since $c(t)$ is continuous for $t \in [0,T]$, it shows that $c(t)$ is a constant for $t \in [0,T]$. □

The Theorem 5.4 talks about Cauchy problem with terminal data for equation (5.8) in the domain $[0,T]\times\mathbb{R}^d$. In the cylindrical domain $Q^{r^2,r}$ with surface $\partial Q^{r^2,r}$ for $r \in (0,1]$, we consider the first boundary problem to the following parabolic equation on $\overline{Q^{r^2,r}}$ with assuming that f is a continuous function on $\partial Q^{r^2,r}$:

$$\begin{cases} \mathcal{L}u(t,x) = D_tu(t,x) + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{d}\partial_i(a_{ij}(t,x)\partial_ju(t,x)) = 0 \qquad on \quad Q^{r^2,r}, \\ u(t,x) = f(t,x) \qquad on \quad \partial Q^{r^2,r}, \end{cases}$$

where $(a_{ij}) = \sigma\sigma^*$. If Assumption 4 (iii) and (iv) hold, from [55, Theorem 3.1] and [55, Corollary 3.2] the solution $u(t,x)$ has a representation as following:

$$u(t,x) = \int_{\partial Q^{r^2,r}} f(s,y)p(s,y;t,x)dS(s,y),$$

where dS denotes the surface measure on $\partial Q^{r^2,r} := ((0,r^2)\times\partial B_r)\cup(\{r^2\}\times B_r)$, and $p(s,y;t,x)$ is the Poisson kernel on $Q^{r^2,r}$ corresponding to the above partial differential equation, which has the following upper bound estimation on $Q^{r^2,r}$ ([55]) with a constant k independent of f

$$p(s,y;t,x) \leqslant k\frac{\exp(-c\frac{|y-x|^2}{s-t})}{(s-t)^{(1+d)/2}} \tag{5.37}$$

for all $(t,x) \in Q^{r^2,r}$, $(s,y) \in \partial Q^{r^2,r}$, $0 \leqslant t < s$.

On the other hand, we can solve the above equation in a probabilitical way. Let

$$\tau_r := \inf\left\{s \geqslant 0 : (s, Y(s,t,x)) \notin Q^{r^2,r}\right\},$$

applying Itô's formula to $u(s,Y(s,t,x))$ and taking expectation, we have for $(t,x) \in Q^{r^2,r}$,

$$u(t,x) = E^{(t,x)}[u(\tau_r, Y(\tau_r,t,x))] - E^{(t,x)}[\int_t^{\tau_r}\mathcal{L}u(s,Y(s,t,x))ds] = E^{(t,x)}[f(\tau_r,Y(\tau_r,t,x))].$$

Hence

$$E^{(t,x)}[f(\tau_r, Y(\tau_r, t, x))] = \int_{\partial Q^{r^2,r}} f(s,y)p(s,y;t,x)dS(s,y).$$

We take $(0,0)$ as the start point of the process $(s, Y(s,t,x))$, then denote $Y_s =: Y(s,0,0)$ and $E[f(\tau_r, Y_{\tau_r})] = \int_{\partial Q^{r^2,r}} f(s,y)p(s,y;0,0)dS(s,y)$.

Lemma 5.10. *If Assumption 4 (iii) and (iv) hold, then on an extension of the probability space there is a stopping time γ such that the distribution of (γ, Y_γ) has a bounded density concentrated on $Q^{1,1}$.*

Proof. Let $n = d+3$. On an extension of our probability space there exists a random variable ρ with values in $[0,1]$ and density function $h(r) = nr^{n-1}$ such that ρ is independent of all $\mathcal{F}_t$. Then ρ is also independent to Y_t, since Y_t is adapted to $\mathcal{F}_t$. Let $\hat{\mathcal{F}}_t = \mathcal{F}_t \vee \sigma(\rho)$, and define γ as the first exit time of (t, Y_t) from $Q^{\rho^2,\rho}$. Then γ is a bounded $\hat{\mathcal{F}}_t$ stopping time. We claim that γ is a random variable of the type that we are looking for.

Actually, according to independence and the above potential knowledge, for a nonnegative continuous function $f(t,x)$ on $[0,\infty)\times\mathbb{R}^d$ we have

$$\begin{aligned}
Ef(\gamma, Y_\gamma) =& E[Ef(\tau_r, Y_{\tau_r})\Big|_{\rho=r}] = E[\int_{\partial Q^{r^2,r}} f(s,y)p(s,y;0,0)dS(s,y)\Big|_{\rho=r}] \\
=& \int_0^1 h(r)dr \int_{\partial Q^{r^2,r}} f(s,y)p(s,y;0,0)dS(s,y) \\
=& \int_0^1 h(r)dr \int_{(0,r^2)\times\partial B_r} f(s,y)p(s,y;0,0)dS(s,y) \\
&+ \int_0^1 h(r)dr \int_{B_r} f(r^2,y)p(r^2,y;0,0)dy =: I_1 + I_2.
\end{aligned}$$

Then (5.37), and the fact that $\exp(-c\frac{|y|^2}{s})s^{-(d+1)/2}$ is bounded by Nr^{-d-1} on $(0,r^2)\times\partial B_r$ yield

$$\begin{aligned}
I_1 &\leqslant k\int_0^1 h(r)dr \int_0^{r^2}\int_{\partial B_r} f(s,y)\frac{\exp(-c\frac{|y|^2}{s})}{s^{(d+1)/2}}dS(s,y) \\
&\leqslant N\int_0^1 h(r)r^{-d-1}dr\int_0^{r^2}\int_{\partial B_r} f(s,y)dS(s,y) \\
&\leqslant N\int_0^1\int_0^{r^2}\int_{\partial B_1} r^{-d-1}f(s,ry)h(r)r^{d-1}d(\partial B_1)dsdr \\
&\leqslant N\int_0^1\int_0^1\int_{\partial B_1} f(s,ry)r^d d(\partial B_1)dsdr \leqslant N\int_{Q^{1,1}} f(s,y)dsdy,
\end{aligned}$$

and

$$
\begin{aligned}
I_2 &\leqslant k\int_0^1\int_{B_r} f(r^2,y)h(r)\frac{\exp(-c\frac{|y|^2}{r})}{r^{d+1}}dydr\\
&\leqslant N\int_0^1\int_{B_r} f(r^2,y)h(r)r^{-d-1}dydr\\
&= N\int_0^1\int_{B_r} f(r^2,y)r^{n-2-d}dydr \leqslant N\int_{Q^{1,1}} f(s,y)dsdy.
\end{aligned}
$$

Hence

$$
Ef(\gamma,Y_\gamma)\leqslant N\int_{Q^{1,1}} f(t,x)dxdt
$$

and N is independent of f.

For arbitrary nonnegative function $fI_{Q^{1,1}}\in\mathbb{L}_1^1$, we can use a standard method to approximate f via continuous functions. The conclusion is proved. □

Lemma 5.11. *Let Assumption 4 hold. Let $K_2\in[0,\infty)$ be a constant. Assume that for some p, q satisfying (5.2) we have*

$$
\psi I_{Q^{1,1}}\leqslant K_2,\quad \|\nabla\psi I_{Q^{1,1}}\|_{\mathbb{L}^q_{\mathbf{p}}}\leqslant K_2.
$$

Take an $r\in(1,\infty)$ and a nonnegative Borel function $f=f(t,x)$ on $(0,\infty)\times\mathbb{R}^d$ such that $f(t,x)=0$ for $t>T$. For $0\leqslant s\leqslant t\leqslant T$ and $x\in\mathbb{R}^d$ introduce

$$
\begin{aligned}
&\rho_t(s,x)=\exp(-\int_s^t\nabla\psi^*\sigma(u,Y(u,s,x))dW_s-\frac{1}{2}\int_s^t|\nabla\psi^*\sigma\sigma^*\nabla\psi|(u,Y(u,s,x))du),\\
&\alpha_t(s,x)=\exp(-2\int_s^t(D_t\psi)_+(u,Y(u,s,x))du),\\
&u_t(s,x)=E\rho_t(s,x)\alpha_t(s,x)f(t,Y(t,s,x)).
\end{aligned}
$$

Then there is a constant N, depending only on K, r, $\mathbf{p}$, q, K_2 and T, such that

$$
\int_0^T u_t(0,0)dt\leqslant N(\int_{(0,\infty)\times\mathbb{R}^d} f^re^{-2\psi}dtdx)^{1/r}+N(\int_{Q^{1,1}} f^{d+3}dtdx)^{1/(d+3)}. \tag{5.38}
$$

Proof. By the strong Markov property of Y, which can be obtained by the same argument as in the proof of Theorem 4.1 which was derived from the strong Feller property of Y_t to SDE (5.32), for any stopping time τ we have

$$
EI_{\tau\leqslant t}\rho_t(0,0)\alpha_t(0,0)f(t,Y_t)=EI_{\tau\leqslant t}\rho_\tau(0,0)\alpha_\tau(0,0)u_t(\tau,Y_\tau).
$$

Therefore, upon assuming without losing generality that $T\geqslant 1$, for γ from Lemma 5.10,

$$
\int_0^T u_t(0,0)dt=E\int_0^\gamma\rho_t(0,0)\alpha_t(0,0)f(t,Y_t)dt+E\rho_\gamma(0,0)\alpha_\gamma(0,0)\int_\gamma^T u_t(\gamma,Y_\gamma)dt=:I_1+I_2.
$$

Observe that $\alpha_t \leqslant 1$ and for $t \leqslant \gamma$ we have $(t, Y_t) \in Q^{1,1}$ so that, in particular, in the formula defining $\rho_t(0,0)$ we can replace $\nabla\psi$ with $\nabla\psi I_{Q^{1,1}}$ and hence all moments of $\rho_t(0,0)I_{t\leqslant\gamma}$ and $\rho_\gamma(0,0)$ are finite and uniformly bounded in t. Since by (5.19) we have

$$E[\exp(\frac{1}{2}\int_s^t |\nabla\psi^*\sigma\sigma^*\nabla\psi| I_{Q^{1,1}}(u, Y(u,s,x))du)] \leqslant C(\|\nabla\psi I_{Q^{1,1}}\|_{\mathbb{L}_p^q}, K) < \infty$$

for all $t \in [0,T]$. For the moments of $\rho_t(0,0)I_{t\leqslant\gamma}$ and $\rho_\gamma(0,0)$, by using the same way of treating (5.27) we get the desired results. With the same argument we can also replace $(\frac{1}{2}\sum_{j=1}^d \int_s^t \partial_j a_{ij}(r, Y(r,s,x))dr)_{1\leqslant i\leqslant d}$ by $(\frac{1}{2}\sum_{j=1}^d \int_s^{t\wedge\gamma} \partial_j a_{ij}(r, Y(r,s,x))dr))_{1\leqslant i\leqslant d}$ in SDE (5.32), it follows by Hölder's inequality and (5.20) that for any $v \in (1,\infty)$

$$I_1 \leqslant N(E\int_0^T f^v(t,Y_t)I_{Q^{1,1}}(t,Y_t)dt)^{1/v} \leqslant N\|f^v I_{Q^{1,1}}\|_{\mathbb{L}^{d+5/2}}^{1/v}.$$

We can choose v so that $v(d+5/2) = d+3$, and get that I_1 is less than the second term on the right in (5.38).

In what concerns I_2 we again use $\alpha_\gamma(0,0) \leqslant 1$ and the finiteness of all moments of $\rho_\gamma(0,0)$. Then we find

$$I_2 \leqslant N(\int_0^1 \int_s^T (\int_{B_1} u_t^r(s,x)dx)dtds)^{1/r}. \tag{5.39}$$

To estimate the interior integral with respect to x we insert there $\exp(-2\psi(s,x))$ and again use Hölder's inequality and the fact that $E\rho_t(s,x) \leqslant 1$. This yields

$$I_2(s,t) := \int_{B_1} u_t^r(s,x)dx \leqslant e^{2K_2}\int_{\mathbb{R}^d} e^{-2\psi(s,x)}\hat{v}_t(s,x)dx$$

where

$$\hat{v}_t(s,x) = E\rho_t(s,x)\alpha_t(s,x)f^r(t,Y(t,s,x)) \leqslant E\beta_t(s,x)f^r(t,Y(t,s,x)).$$

Hence by Lemma 5.9,

$$I_2(s,t) \leqslant e^{2K_2}\int_{\mathbb{R}^d} e^{-2\psi(t,x)}f^r(t,x)dx,$$

which shows that I_2 is less than the first term on the right in (5.38). The Lemma is proved. □

Lemma 5.12. *Let the assumptions of Lemma 5.11 be satisfied and let $\epsilon \in [0,2)$ be a constant and h a nonnegative Borel function on bounded domain $Q \subset \mathbb{R}^{d+1}$ such that on Q,*

$$2D_t\psi + \sum_{i,j=1}^d \partial_j(a_{ij}\partial_i\psi) \leqslant he^{\epsilon\psi}. \tag{5.40}$$

Then for any $\delta \in [0, 2-\epsilon)$, $r \in (1, 2/(\delta+\epsilon)]$, there exists a constant N, depending only on K, T, p, q, K_2, ϵ, δ and r (but not Q) such that for any stopping time $\tau \leqslant \tau_Q(Y.)$ we have

$$E\Phi_\tau \leqslant N + N\Big(\int_Q h^r e^{-(2-r\eta)\psi}\,dtdx\Big)^{1/r} + N\sup_{Q^{1,1}} h, \tag{5.41}$$

where $\eta = \delta + \epsilon$ so that $r\eta \leqslant 2$ and

$$\begin{aligned}\Phi_t := \exp\Big(&- \int_0^t (\nabla\psi^*\sigma)(s, Y_s)dW_s - \frac{1}{2}\int_0^t |\nabla\psi^*\sigma\sigma^*\nabla\psi|(s, Y_s)ds \\ &- 2\int_0^t (D_t\psi)_+(s, Y_s)ds + \delta\psi(t, Y_t)\Big).\end{aligned}$$

Proof. By Itô's formula,

$$\begin{aligned}\Phi_\tau = &\Phi_0 + m_\tau + \int_0^\tau \Phi_t\Big[\delta D_t\psi + \frac{\delta}{2}\sum_{i,j=1}^d \partial_j(a_{ij}\partial_i\psi) - 2(D_t\psi)_+ \\ &+ \frac{1}{2}(|\delta-1|^2 - 1)|\nabla\psi^*\sigma\sigma^*\nabla\psi|\Big](t, Y_t)dt\end{aligned}$$

where m_t is a local martingale starting at zero. By using (5.40), and the inequality $|\delta - 1| \leqslant 1$ we obtain

$$\Phi_\tau \leqslant \Phi_0 + \delta\int_0^\tau \Phi_t h(t, Y_t)\exp(\epsilon\psi(t, Y_t))dt + m_\tau. \tag{5.42}$$

Since $\Phi_t \geqslant 0$ we take the expectations of both sides and drop Em_τ. More precisely, we introduce $\tau_n := \inf\{t \geqslant 0 : |m_t| \geqslant n\}$ and substitute $\tau \wedge \tau_n$ in place of τ in (5.42). After that we take expectations, use the fact that $Em_{\tau\wedge\tau_n} = 0$, let $n \to \infty$, and finally use Fatou's Lemma with monotone convergence theorem. Furthermore, we denote $f = I_Q h\exp(\eta\psi)$ and notice that $\tau \leqslant T$. Then in the notation of Lemma 5.11, we find that

$$\begin{aligned}E\Phi_\tau &\leqslant N + NE\int_0^\tau \rho_t(0,0)\alpha_t(0,0)f(t, Y_t)dt \\ &\leqslant N + N\int_0^T E\rho_t(0,0)\alpha_t(0,0)f(t, Y_t)dt = N + N\int_0^T u_t(0,0)dt.\end{aligned}$$

It only remains to note that the first term in the right-hand side of (5.38) is just the second one on the right in (5.41) and the second integral on the right in (5.38) is less than $volQ^{1,1}\sup_{Q^{1,1}} h^{d+3}\exp[\eta K_2(d+3)]$. The Lemma is proved. □

Theorem 5.13. *Let Assumption 4 hold. Let K_1, $K_2 \in [0, \infty)$ and $\epsilon \in [0, 2)$ be some constants and let Q be a bounded subdomain of Q_T and h be a nonnegative Borel function on Q. Assume that for some $\mathbf{p}$, q satisfying (5.2) we have*

$$hI_{Q^{1,1}} \leqslant K_2, \quad \psi I_{Q^{1,1}} \leqslant K_2, \quad \|\nabla\psi I_{Q^{1,1}}\|_{\mathbb{L}^q_{\mathbf{p}}} \leqslant K_2.$$

Also assume that on Q

$$\psi \geqslant 0, \quad 2D_t\psi \leqslant K_1\psi,$$
$$2D_t\psi + \sum_{i,j=1}^{d} \partial_j(a_{ij}\partial_i\psi) \leqslant he^{\epsilon\psi}.$$

Denote by X_t, $t \in [0,T]$, *the solution of*

$$X_t = \int_0^t \sigma(s, X_s)dW_s + \int_0^t (-\sigma\sigma^*\nabla\psi)(s, X_s)ds + (\frac{1}{2}\sum_{j=1}^{d}\int_0^t \partial_j a_{ij}(s, X_s)ds)_{1\leqslant i\leqslant d}.$$

Then for any $r \in (1, 4/(2+\epsilon)]$ *there exists a constants* N, *depending only on* K, K_1, K_2, r, d, T, $\mathbf{p}$, q, *and* ϵ, *such that*

$$E \sup_{t\leqslant\tau_Q(X.)} \exp[\mu(\psi(t, X_t) + \nu|X_t|^2)] \leqslant N + NH_Q(T, a, r) \tag{5.43}$$

where H_Q *is introduced in Assumption 1,* $a = (2 - r\eta)\nu$, $\eta = 2\delta + \epsilon$, μ, ν *and* δ *are taken from* (5.6). *Here* $\tau_Q(X.) := \inf\{t \geqslant 0 : (t, X_t) \notin Q\}$.

Proof. Define $\hat\psi = \psi + \nu|x|^2$,

$$M_t = \exp(\delta\hat\psi(t, X_t) - \frac{K_1}{2}\int_0^t \hat\psi(s, X_s)ds), \quad M_* = \sup_{t\leqslant\tau_Q(X.)} M_t.$$

Then for $t \leqslant \tau_Q(X.)$,

$$\hat\psi(t, X_t) \leqslant \ln M_*^{1/\delta} + \frac{K_1}{2\delta}\int_0^t \hat\psi(s, X_s)ds$$

and hence by Gronwall's inequality

$$\hat\psi(t, X_t) \leqslant e^{tK_1/(2\delta)} \ln M_*^{1/\delta} \leqslant e^{TK_1/(2\delta)} \ln M_*^{1/\delta}.$$

Take $\mu = \frac{\delta}{2}e^{-TK_1/(2\delta)}$, then

$$\exp(\mu\hat\psi(t, X_t)) \leqslant \sqrt{M_*}. \tag{5.44}$$

Therefore, to prove (5.43), it suffices to prove that $E\sqrt{M_*} \leqslant N$. It turns by a well known result on transformations of stochastic inequalities (see Lemma 3.2 in [30]), if $EM_\tau \leqslant N_1$ for all stopping times $\tau \leqslant \tau_Q(X.)$. Then $E\sqrt{M_*} \leqslant 3N_1$. Thus, it suffices to estimate EM_τ.

On a probability space carrying a $d-$dimensional Wiener process $\hat W_t$ introduce $\hat X_t$ as the solution of the equation

$$\hat X_t = \int_0^t \sigma(s, \hat X_s)d\hat W_s - \int_0^{t\wedge\tau_Q(\hat X.)} \sigma\sigma^*\nabla\hat\psi(s, \hat X_s)ds + (\int_0^{t\wedge\tau_Q(\hat X.)} \frac{1}{2}\sum_{j=1}^{d} \partial_j a_{ij})(s, \hat X_s)ds)_{1\leqslant i\leqslant d}. \tag{5.45}$$

Also set

$$\hat{M}_t = \exp(2\delta\hat{\psi}(t,\hat{X}_t) - 2\int_0^t (D_t\hat{\psi})_+(s,\hat{X}_s)ds).$$

Write $\hat{E}$ for the expectation sign on the new probability space and observe that on Q

$$2D_t\hat{\psi} + \sum_{i,j=1}^d \partial_j(a_{ij}\partial_i\hat{\psi}) = 2D_t\psi + \sum_{i,j=1}^d \partial_j(a_{ij}\partial_i\psi) + 2\nu\sum_{i,j=1}^d x_i\partial_j a_{ij} + 2\nu\sum_{i,j=1}^d \partial_j a_{ij}$$
$$\leqslant (h+C)e^{\epsilon\hat{\psi}}. \tag{5.46}$$

Here $2\nu\sum_{i,j=1}^d x_i\partial_j a_{ij} + 2\nu\sum_{i,j=1}^d \partial_j a_{ij} \leqslant (h+C)e^{\epsilon\hat{\psi}}$ holds because of Assumption 4, which means $|\partial_j a_{ij}|$ is bounded. Then after an obvious change of measure (cf. Lemma A.3) inequality (5.41) with 2δ, $\hat{E}$, $\hat{\psi}$, and $\hat{W}_t$ in place of δ, E, ψ, and W_t, respectively, $\eta = 2\delta+\epsilon$, and $r\in(1,4/(2+\epsilon)]\subset(1,2/(2\delta+\epsilon)]$ is written as

$$\hat{E}\hat{M}_\tau \leqslant N + N(\int_Q h^r I_{(0,T)}e^{-(2-r\eta)\hat{\psi}}dtdx)^{1/r}$$

and since $\hat{\psi}\geqslant\nu|x|^2$ on Q, we obtain

$$\hat{E}\hat{M}_\tau \leqslant N + N(\int_Q h^r I_{(0,T)}e^{-(2-r\eta)\nu|x|^2}dtdx)^{1/r} = N + NH_Q^{1/r}(T,(2-r\eta)\nu,r) =: N_0$$

for all stopping times $\tau\leqslant\tau_Q(\hat{X}_\cdot)$, which yields

$$\hat{E}\sqrt{\hat{M}_*}\leqslant 3N_0.$$

Combining this with the inequality

$$\exp(2\delta\hat{\psi}(t,\hat{X}_t) - K_1\int_0^t \hat{\psi}(x,\hat{X}_s)ds) \leqslant \hat{M}_t, \qquad t\leqslant\tau_Q(\hat{X}_\cdot),$$

the left-hand side of which is quite similar to M_t but with $2\hat{\psi}$ in place of $\hat{\psi}$, the above argument deduce

$$\hat{E}\sup_{t\leqslant\tau_Q(\hat{X}_\cdot)}\exp(2\mu\nu|\hat{X}_t|^2) \leqslant \hat{E}\sup_{t\leqslant\tau_Q(\hat{X}_\cdot)}\exp(2\mu\hat{\psi}(t,\hat{X}_t)) \leqslant NN_0. \tag{5.47}$$

We now estimate EM_τ through $\hat{E}\hat{M}_\tau$ by using Girsanov's theorem and Hölder's inequality. We use a certain freedom in choosing $\hat{X}_t$ and $\hat{W}_t$ and on the probability space where W_t and X_t are given we introduce a new measure by the formula:

$$\hat{P}(d\omega) = \exp(-2\nu\int_0^\infty X_t^*\sigma(t,X_t)I_{t<\tau_Q(X_\cdot)}dW_t - 2\nu^2\int_0^\infty X_t^*(\sigma\sigma^*)(t,X_t)X_tI_{t<\tau_Q(X_\cdot)}dt)P(d\omega).$$

Since Q is a bounded domain, then we have

$$E\exp\Big(2\nu^2\int_0^\infty X_t^*(\sigma\sigma^*)(t,X_t)X_tI_{t<\tau_Q(X.)}dt\Big)\leqslant E\exp\Big(2\nu^2K\int_0^T X_t^*X_tI_{t<\tau_Q(X.)}dt\Big)<\infty,$$

which implies that $\hat{P}$ is a probability measure. Furthermore, as is easy to see, for $t\leqslant\tau_Q(X.)$

$$\hat{X}_t:=X_tI_{t<\tau_Q(X.)}+(\int_0^t\sigma(s,X_s)dW_s-\int_0^{\tau_Q(X.)}\sigma(s,X_s)dW_s+X_{\tau_Q(X.)})I_{t\geqslant\tau_Q(X.)}$$

coincides with X_t and satisfies (5.45) for $t\leqslant\tau_Q(X.)$ with

$$\hat{W}_t=W_t+2\nu\int_0^{t\wedge\tau_Q(X.)}\sigma^*(s,X_s)X_sds$$

which is a Wiener process with respect to $\hat{P}$. In this situation for $\tau\leqslant\tau_Q(X.)=\tau_Q(\hat{X}.)$

$$\begin{aligned}EM_\tau&\leqslant\hat{E}\hat{M}_\tau^{1/2}\exp(2\nu\int_0^\infty\hat{X}_t^*\sigma(t,\hat{X}_t)I_{t<\tau_Q(\hat{X}.)}d\hat{W}_t-2\nu^2\int_0^\infty\hat{X}_t^*(\sigma\sigma^*)(t,\hat{X}_t)\hat{X}_tI_{t<\tau_Q(\hat{X}.)}dt)\\&\leqslant(\hat{E}\hat{M}_\tau)^{1/2}(\hat{E}\rho^{1/2}\exp(12v^2\int_0^\infty\hat{X}_t^*(\sigma\sigma^*)(t,\hat{X}_t)\hat{X}_tI_{t<\tau_Q(\hat{X}.)}dt))^{1/2}\end{aligned}$$

where

$$\rho=\exp(8\nu\int_0^\infty\hat{X}_t^*\sigma(t,\hat{X}_t)I_{t<\tau_Q(\hat{X}.)}d\hat{W}_t-32\nu^2\int_0^\infty\hat{X}_t^*(\sigma\sigma^*)(t,\hat{X}_t)\hat{X}_tI_{t<\tau_Q(\hat{X}.)}dt).$$

Observe that $\hat{E}\rho=1$ and $\hat{E}\hat{M}_\tau\leqslant N_0$. Therefore,

$$EM_\tau\leqslant N_0^{1/2}(\hat{E}\exp(24\nu^2\int_0^{\tau_Q(X.)}(\hat{X}_t^*(\sigma\sigma^*)(t,\hat{X}_t)\hat{X}_t)dt))^{1/4}.$$

It only remains to refer to (5.47) after noticing that

$$24\nu^2\int_0^{\tau_Q(X.)}(\hat{X}_t^*(\sigma\sigma^*)(t,\hat{X}_t)\hat{X}_t)dt\leqslant24\nu^2KT\sup_{t\leqslant\tau_Q(X.)}|X_t|^2=2\mu\nu\sup_{t\leqslant\tau_Q(X.)}|X_t|^2$$

and use the inequality $\iota^\alpha\leqslant1+\iota$ if $\iota\geqslant0$, $0\leqslant\alpha\leqslant1$, where $\nu=\mu/(12KT)$. The theorem is proved. □

5.4 Proof of Theorem 5.2

By Theorem 4.1 the strong solution X_t to (5.5) is defined at least until the time ξ when $(s+t,X_t)$ exits from all Q^n. We claim that in order to prove $\xi=\infty$ $(a.s.)$ and also to prove the second assertions of the theorem, it suffices to prove that for each $T\in(0,\infty)$ and $m\geqslant1$ there exists a constant N, depending only on K, K_1, d, $\mathbf{p}(m+1)$, $q(m+1)$, ϵ,

T, $\|\nabla\phi I_{Q^{m+1}}\|_{\mathbb{L}^{q(m+1)}_{\mathbf{p}(m+1)}}$, $dist(\partial Q^m, \partial Q^{m+1})$, $\sup_{Q^{m+1}}\{\phi+h\}$, M and the function H, such that for $(s,x)\in Q^m$ we have

$$E\sup_{t<\xi\wedge T}\exp(\mu\phi(s+t,X_t)+\mu\nu|X_t|^2)\leqslant N. \tag{5.48}$$

To prove the claim notice that (5.48) implies

$$\sup_{t<\xi\wedge T}(\phi(s+t,X_t)+|X_t|^2)<\infty \quad (a.s.). \tag{5.49}$$

It follows that $(a.s.)$ there exists an $n\geqslant 1$ such that up to time $\xi\wedge T$ the trajectory $Z_t=(s+t,X_t)$ lies in Q^n. Indeed, on the set of all ω where this is wrong, for the exit time ξ^n of Z_t from Q^n we have $\xi^n<T$ for all n. However owing to (5.49), the sequence X_{ξ^n} should be bounded, then the sequence Z_{ξ^n} has limit points on the boundary ∂Q. According to the Assumption 1 (vi), it only happens with probability zero. Hence, $(a.s.)$ there is $n\geqslant 1$ such that $T\leqslant\xi^n$. Since this happens for any T and $\xi^n<\xi$ we conclude that $\xi=\infty$ $(a.s.)$, which proves our intermediate claim.

Since $dist(\partial Q^m,\partial Q^{m+1})>0$ we can find $\kappa\in(0,1]$ sufficiently small so that $(s,x)+Q^{\kappa^2,\kappa}\subset Q^{m+1}$ for all $(s,x)\in Q^m$. Therefore, by translation and dilation, without losing generality, we may assume that $s=0$, $x=0$ and $Q^{1,1}\subset Q^m$.

Next we notice that obviously, to prove (5.48) it suffices to prove that with N of the same kind as in (5.48) for any $n\geqslant m+2$,

$$E\sup_{t<\xi^n\wedge T}\exp(\mu\phi(t,X_t)+\mu\nu|X_t|^2)\leqslant N. \tag{5.50}$$

Fix an $n\geqslant m+2$. By virtue of Theorem 4.1, notice that the left-hand side of (5.50) will not change if we change $-\sigma\sigma^*\nabla\phi+(\frac{1}{2}\sum_{j=1}^d\partial_j a_{ij})_{1\leqslant i\leqslant d}$ outside of Q^n. Therefore we may replace ϕ with $\phi\eta$ and replace $\frac{1}{2}\sum_{j=1}^d\partial_j a_{ij}$ with $\frac{1}{2}\sum_{j=1}^d\partial_j a_{ij}\eta$ for each $1\leqslant i\leqslant d$, where η is an infinitely differentiable function equal 1 on a neighborhood of Q^n and equals 0 outside of Q^{n+1}. To simplify the notation we just assume that ϕ and $\frac{1}{2}\sum_{j=1}^d\partial_j a_{ij}$ vanishes outside of Q^{n+1} and (5.3) holds in a neighborhood of $\overline{Q^n}$. This is harmless as long as we prove that N depends appropriately on the data.

Now we mollify ϕ by convolving it with a $\delta-$like nonnegative smooth function $\zeta^\gamma(t,x)=\gamma^{-d-1}\zeta(t/\gamma,x/\gamma)$, ζ has compact support in Q^1. Denote by $\phi^{(\gamma)}$ the result of the convolution and use an analogous notation for the convolution of $\zeta^\gamma(t,x)$ with other functions. Also denote by X^γ_t the solution of the following SDE

$$X^\gamma_t=\int_0^t\sigma(s,X^\gamma_s)dW_s+\int_0^t(-\sigma\sigma^*\nabla\phi^{(\gamma)})(s,X^\gamma_s)ds+(\frac{1}{2}\sum_{j=1}^d\int_0^t\partial_j a_{ij}(s,X^\gamma_s)ds)_{1\leqslant i\leqslant d}.$$

For $x.\in\mathcal{C}([0,\infty),\mathbb{R}^d)$ we define $\xi_n(x.):=\inf\{t\geqslant 0:(t,x_t)\notin Q^n\}$. Consider the bounded function f on $\mathcal{C}([0,\infty),\mathbb{R}^d)$ given by the formula

$$f(x.)=\sup_{t\leqslant\xi^n(x.)\wedge T}\exp(\mu\phi(t,x_t)+\mu\nu|x_t|^2),$$

and let f^γ be defined by the same formula with $\phi^{(\gamma)}$ in place of ϕ. Since $\sigma\sigma^*$ is bounded, by using Lemma 5.8 we conclude that the left-hand side of (5.50) is equal to the limit as $\gamma \downarrow 0$ of

$$Ef^\gamma(X^\gamma_\cdot) = E \sup_{t<\xi^n(X^\gamma_\cdot)\wedge T} \exp(\mu\phi^{(\gamma)}(t, X^\gamma_t) + \mu\nu|X^\gamma_t|^2). \tag{5.51}$$

In fact, if we denote $M_t = \int_0^t \sigma(s, M_s)dW_s$, according to Lemma 5.8

$$\begin{aligned}|Ef(X_\cdot) - Ef^\gamma(X^\gamma_\cdot)| &\leqslant N'(E|f(M_\cdot) - f^\gamma(M_\cdot)|^2)^{1/2} + N'\|f\|_\infty\|\sigma\sigma^*(\nabla\phi - \nabla\phi^{(\gamma)})I_{Q^n}\|_{L^q_{\mathbf{p}}} \\ &\leqslant N'(E|f(M_\cdot) - f^\gamma(M_\cdot)|^2)^{1/2} + KN'\|(\nabla\phi - \nabla\phi^{(\gamma)})I_{Q^n}\|_{L^q_{\mathbf{p}}},\end{aligned}$$

which of course tends to 0 when $\gamma \to 0$, since ϕ is continuous and bounded on Q^n, $\nabla\phi I_{Q^n} \in \mathbb{L}^q_{\mathbf{p}}$, then $f^\gamma \to f$ and $\nabla\phi^{(\gamma)} I_{Q^n} \to \nabla\phi I_{Q^n}$ in $\mathbb{L}^q_{\mathbf{p}}$ as $\gamma \to 0$.

In the light of the fact that (5.3) holds in a neighborhood of Q^n we have that on Q^n for sufficiently small γ

$$\begin{aligned}2D_t\phi^{(\gamma)} + \sum_{i,j=1}^d \partial_j(a_{ij}\partial_i\phi^{(\gamma)}) &\leqslant ((he^{\epsilon\phi})^{(\gamma)}e^{-\epsilon\phi^{(\gamma)}} + \sum_{i,j=1}^d |\partial_j(a_{ij}\partial_i\phi^{(\gamma)}) - (\partial_j(a_{ij}\partial_i\phi))^{(\gamma)}|)e^{\epsilon\phi^{(\gamma)}} \\ &=: h^\gamma e^{\epsilon\phi^{(\gamma)}}.\end{aligned} \tag{5.52}$$

Since h is continuous, then $(he^{\epsilon\phi})^{(\gamma)}e^{-\epsilon\phi^{(\gamma)}} \to h$ uniformly on Q^n. Besides $\sum_{i,j=1}^d |\partial_j(a_{ij}\partial_i\phi^{(\gamma)}) - (\partial_j(a_{ij}\partial_i\phi))^{(\gamma)}|) \to 0$ pointwise. Hence if we denote

$$H^\gamma_{Q^n}(T, (2-r\eta)\nu, r) := \int_{Q^n} (h^\gamma)^r(t,x)I_{(0,T)}(t)e^{-(2-r\eta)\nu|x|^2}dtdx,$$

we have

$$\lim_{\gamma\to 0} H^\gamma_{Q^n}(T, (2-r\eta)\nu, r) \leqslant H_{Q^n}(T, (2-r\eta)\nu, r).$$

Furthermore, the conditions $2D_t\phi^{(\gamma)} \leqslant K_1\phi^{(\gamma)}$ also hold in a neighborhood of Q^n for sufficiently small γ.

We now apply Theorem 5.13 for $Q^n \cap Q_T$ in place of Q to conclude that

$$\begin{aligned}E\sup_{t<\xi^n\wedge T} \exp(\mu\phi(t,X_t) + \mu\nu|X_t|^2) &= \lim_{\gamma\downarrow 0} E \sup_{t<\xi^n(X^\gamma_\cdot)\wedge T} \exp(\mu\phi^{(\gamma)} + \mu\nu|X^\gamma_t|^2) \\ &\leqslant \lim_{\gamma\downarrow 0}(N + NH^\gamma_{Q^n}(T, (2-r\eta)\nu, r)) \\ &\leqslant N + NH_{Q^n}(T, (2-r\eta)\nu, r) \\ &\leqslant N + NH_Q(T, (2-r\eta)\nu, r),\end{aligned}$$

where the values of all the parameters are specified in 5.13 and the constants N depend only on r, d, $\mathbf{p}(m+1)$, $q(m+1)$, ϵ, T, K, K_1, $\|\nabla\phi I_{Q^{m+1}}\|_{L^{q(m+1)}_{\mathbf{p}(m+1)}}$, and $\sup_{Q^{m+1}}\{\phi + h\}$.

We finally use condition (H) from Assumption 1 . Fix any $r_0 \in (1, 2/(2\delta + \epsilon))$, set $a = (2 - r_0\eta)\nu$ (> 0) and take $r = r(T, a)$ from condition (H). Hölder's inequality shows that if condition (H) is satisfied with $r = r'$ where $r' > 1$, then it is also satisfied with any $r \in (1, r']$. Hence without losing generality we may assume that $r = r(T, a) \in (1, r_0]$. Then $(2 - r\eta)\nu \geqslant a$ and $H_Q(T, (2 - r\eta)\nu, r) \leqslant H_Q(T, a, r(T, a)) < \infty$. Thus, Theorem 5.13 yields (5.50). The theorem is proved.

□

Remark 5.14. We can add another drift term to (5.5), it does not have to be the gradient of a function. Under Assumption 1 take a Borel measurable locally bounded $\mathbb{R}^d$ valued function $b(t,x)$ defined on $\mathbb{R}^{d+1}$ satisfying the condition $|b(t,x)| \leqslant c(1+|x|)$, where c is a finite positive constant, then it turns out that the first assertion of Theorem 5.2 still holds with the equation

$$X_t = x + \int_0^t \sigma(s+r, X_r)dW_r + \int_0^t (-\sigma\sigma^*\nabla\phi)(s+r, X_r)dr + \int_0^t b(s+r, X_r)dr + \Big(\int_0^t \frac{1}{2}\sum_{j=1}^d \partial_j a_{ij}(s+r, X_r)dr\Big)_{1\leqslant i\leqslant d} \tag{5.53}$$

in place of (5.5). To prove this we follow the proof in [37] Remark 8.2. The only needed material is the Markov property of solution to equation (5.5), which we already get from the proof of Theorem 4.1. By applying Girsanov theorem we get the non-explosion result for the equation (5.53).

Further we can carry our results in Theorem 5.2 to the cases in which ϕ is not necessarily nonnegative but $\phi \geqslant -C(1+|x|^2)$, $C > 0$. Since the equation (5.5) is equivalent to the following

$$X_t = x + \int_0^t \sigma(s+r, X_r)dW_r + \Big(\frac{1}{2}\int_0^t \sum_{j=1}^d \partial_j a_{ij}(s+r, X_r)dr\Big)_{1\leqslant i\leqslant d} + \int_0^t 2C\sigma\sigma^*(s+r, X_r)X_r dr - \int_0^t \sigma\sigma^*\nabla[C(1+|x|^2)+\phi](s+r, X_r)dr,$$

obviously $|\sigma\sigma^*(t,x)x| \leqslant K(1+|x|)$. We conclude that SDE (5.5) has a unique solution defined for all times if $(s,x) \in Q$ provided that $\phi + C(1+|x|^2)$ rather than ϕ satisfies Assumption 1.

5.5 Diffusions in random media

We apply our results to a particle which performs a random motion in $\mathbb{R}^d$, $d \geqslant 2$, interacting with impurities which are randomly distributed according to a Gibbs measure of Ruelle type. So, the impurities form a locally finite subset $\gamma = \{x_k | k \in \mathbb{N}\} \subset \mathbb{R}^d$. The interaction is given by a pair potential V and diffusion coefficient σ to be specified below defined on $\{x \in \mathbb{R}^d : |x| > \rho\}$, where $\rho \geqslant 0$ is a given constant. The stochastic dynamics of the particle is then determined by a stochastic equation type (5.5) as in Theorem 5.2 above with

$$Q := \mathbb{R}_+ \times (\mathbb{R}^d \backslash \gamma^\rho), \quad \phi(t,x) := \sum_{y\in\gamma} V(x-y), \quad (t,x) \in Q, \tag{5.54}$$

where γ^ρ is the closed $\rho-$neighborhood of the set γ, i.e., the random path X_t of the particle should be the strong solution of

$$X_t = x + \int_0^t \sigma(X_s)dW_s + (\frac{1}{2}\sum_{j=1}^d \int_0^t \partial_j a_{ij}(X_s)ds)_{1\leqslant i\leqslant d} - \sum_{w\in\gamma}\int_0^t (\sigma\sigma^*)(X_s)\nabla V(X_s - w)ds. \tag{5.55}$$

Below we shall give conditions on the pair potential V and diffusion coefficient σ which imply that this is indeed the case, i.e. that Theorem 5.2 above applies, for all γ outside a set of measure zero for the Gibbs measure. Here the original case is from [37] section 9.1, we generalize it to the multiplicative noise case. Similarly the set of admissible impurities γ we can treat is

$$\Gamma_{ad} := \left\{\gamma \subset \mathbb{R}^d | \forall r > 0 \exists c(\gamma,r) > 0 : |\gamma \cap B_r(x)| \leqslant c(\gamma,r)\log(1+|x|), \forall x \in \mathbb{R}^d\right\}, \tag{5.56}$$

where $B_r(x)$ denotes the open ball with center x and radius r, $|A|$ denotes the cardinality of a set A. From [37] we know that for essentially all classes of Gibbs measure in equilibrium statistical mechanics of interacting infinite particle systems in $\mathbb{R}^d$ the set Γ_{ad} has measure one, this is also true for Ruelle measures.

We fix a $\gamma \in \Gamma_{ad}$. The necessary conditions on the pair potential V and diffusion coefficient σ go as follows (the typical case when $\rho = 0$ is also included):

(V1) The function V is positive and once continuously differentiable in $\mathbb{R}^d \cap \{|x| > \rho\}$, $\lim_{|x|\downarrow\rho} V(x) = \infty$.

(V2) There exist finite constants $\alpha > d/2$, $C \geqslant 0$, $\epsilon \in [1,2)$ such that with $U(x) =: C(1+|x|^2)^{-\alpha}$ we have

$$|V(x)| + |\nabla V(x)| \leqslant U(x) \quad \text{for} \quad |x| > \rho, \tag{5.57}$$

and for any $|y| > \rho$

$$\sum_{i,j=1}^d (\partial_j a_{ij}(x)\partial_i V(y) + a_{ij}(x)\partial_i\partial_j V(y)) \leqslant C(e^{\epsilon(V+U)(y)} - 1) \tag{5.58}$$

in the sense of distributions on $\{x \in \mathbb{R}^d : |x| > \rho\}$ where $\sigma(x) = (\sigma_{ij}(x))_{1\leqslant i,j\leqslant d} : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ satisfies the following conditions:

(σ1) There exists a positive constant K such that for all $x \in \mathbb{R}^d$

$$\frac{1}{K}|\lambda|^2 \leqslant \langle(\sigma\sigma^*)(x)\lambda, \lambda\rangle \leqslant K|\lambda|^2, \quad \forall\lambda \in \mathbb{R}^d. \tag{5.59}$$

($\sigma 2$) For $1 \leqslant i,\, j \leqslant d$, $\sigma_{ij} \in \mathcal{C}_b^2(\mathbb{R}^d)$.

We emphasize that above conditions are fulfilled for essentially all potentials of interests in statistical physics.

Introduce $\bar{V}(x) = V(x) + 2U(x)$, $|x| > \rho$, and for $(t, x) \in Q$ let

$$\bar{\phi}(t,x) := \sum_{y\in\gamma} \bar{V}(x-y), \quad (a_{ij})_{1\leqslant i,j\leqslant d} := \sigma\sigma^*,$$
$$b(t,x) := 2\sum_{w\in\gamma}(\sigma\sigma^*)(x)\nabla U(x-w).$$

Owing to (5.57), (5.59) and the fact that $\gamma \in \Gamma_{ad}$, the function ϕ is continuously differentiable in Q and $|b(t,x)| \leqslant NK\log(2+|x|)$, where N is independent of (t,x) (See [37] Section 9.1). Meanwhile for appropriate constants N on Q we have for $|y| > \rho$

$$2\sum_{i,j=1}^{d}(\partial_j a_{ij}(x)\partial_i U(y) + a_{ij}(x)\partial_j\partial_i U(y)) \leqslant N(e^{\epsilon U(y)} - 1)$$

because of conditions ($\sigma 1$) and ($\sigma 2$). Combing this with the fact that $V+U$ is positive and $\sum(e^{a_k}-1) \leqslant e^{\sum a_k} - 1$, $a_k \geqslant 0$, we find that there exists a constant $N' > 0$ independent of (t,x) such that

$$\begin{aligned}\sum_{i,j=1}^{d}\partial_j(a_{ij}\partial_i\bar{\phi})(x) &= \sum_{i,j=1}^{d}\sum_{w\in\gamma}\partial_j(a_{ij}(x)\partial_i(V(x-w)+2U(x-w)))\\ &\leqslant N\sum_{w\in\gamma}\Big((e^{\epsilon(V(x-w)+2U(x-w))}-1)+(e^{\epsilon U(x-w)}-1)\Big) \leqslant N'(e^{\epsilon\bar{\phi}(x)}-1).\end{aligned}$$

It shows that all conditions on $\bar{\phi}$ and σ in Theorem 5.2 are fulfilled and therefore by Remark 5.14 the equation

$$X_t = x + \int_0^t \sigma(X_s)dW_s - \int_0^t (\sigma\sigma^*\nabla\bar{\phi})(X_s)ds + (\frac{1}{2}\sum_{j=1}^{d}\int_0^t \partial_j a_{ij}(X_s)ds)_{1\leqslant i\leqslant d} + \int_0^t b(X_s)ds \tag{5.60}$$

has a unique strong solution defined for all times if $x \in \mathbb{R}^d\backslash\gamma^\rho$. Since equation (5.60) coincides with SDE (5.55), we get the desired conclusion.

5.6 M-particle systems with gradient dynamics

In this subsection we consider a model of M particles in $\mathbb{R}^d$ interacting via a pair potential V and diffusion coefficient σ satisfying the following conditions:

(V1) The function V is once continuously differentiable in $\mathbb{R}^d\backslash\{0\}$, $\lim_{|x|\to 0} V(x) = \infty$,

and on $\mathbb{R}^d\setminus\{0\}$ we assume that $V \geqslant -U$, where $U(x) := C(1+|x|^2)$, C is a constant.
(V2) There exists a constant $\epsilon \in [1,2)$ such that for arbitrary $x, y \in \mathbb{R}^d\setminus\{0\}$,

$$\sum_{i,j=1}^{d}(\partial_j a_{i,j}(x)\partial_i V(y) + a_{i,j}(x)\partial_i\partial_j V(y)) \leqslant Ce^{\epsilon(V+U)(y)} \tag{5.61}$$

in the sense of distributions.
Here $(a_{i,j})_{1\leqslant i,j\leqslant d} := \sigma\sigma^*$ and $\sigma(x) = (\sigma_{i,j}(x))_{1\leqslant i,j\leqslant d} : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ is the diffusion coefficient satisfying:
(σ1) There exists a positive constant K such that for all $x \in \mathbb{R}^d$

$$\frac{1}{K}|\lambda|^2 \leqslant \langle(\sigma\sigma^*)(x)\lambda, \lambda\rangle \leqslant K|\lambda|^2, \quad \forall\lambda \in \mathbb{R}^d,$$

(σ2) For $1 \leqslant i,\ j \leqslant d$, $\sigma_{i,j} \in \mathcal{C}_b^2(\mathbb{R}^d)$.
Introduce $\bar{V} := V + 2U$,

$$Q := \mathbb{R}_+ \times \left(\mathbb{R}^{Md}\setminus \cup_{1\leqslant k<j\leqslant M}\left\{x = (x^{(1)},...,x^{(M)}) \in \mathbb{R}^{Md} : x^{(k)} = x^{(j)}\right\}\right),$$

$$Q^n := (0,n) \times \left\{x = (x^{(1)},...,x^{(M)}) \in \mathbb{R}^{Md} : |x| < n, x^{(k)} \neq x^{(j)} \text{ for } 1 \leqslant k < j \leqslant M\right\},$$

and let the function ϕ, $\bar{\phi}$, $\bar{\sigma}$, $\bar{a}$ and b be defined on Q by

$$\phi(t,x) := \sum_{1\leqslant k<j\leqslant M} V(x^{(k)} - x^{(j)}), \quad \bar{\phi}(t,x) := \sum_{1\leqslant k<j\leqslant M} \bar{V}(x^{(k)} - x^{(j)}),$$

$$\bar{\sigma}(t,x) := \begin{bmatrix} \sigma(x^{(1)}) & 0 & 0 \\ 0 & \sigma(x^{(2)}) & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & \sigma(x^{(M)}) \end{bmatrix}, \bar{a}(t,x) := \begin{bmatrix} (\sigma\sigma^*)(x^{(1)}) & 0 & 0 \\ 0 & (\sigma\sigma^*)(x^{(2)}) & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & (\sigma\sigma^*)(x^{(M)}) \end{bmatrix},$$

$$b := (b^{(1)},...,b^{(M)}), \quad b^{(k)}(t,x) := 4C(\sigma\sigma^*)(x^{(k)}) \sum_{1\leqslant j\neq k\leqslant M} (x^{(k)} - x^{(j)}), \quad k = 1,\cdots,M.$$

Observe that for arbitrary y, $x \in \mathbb{R}^d\setminus\{0\}$,

$$2\sum_{i,j=1}^{d}(\partial_j a_{i,j}(x)\partial_i U(y) + a_{i,j}(x)\partial_j\partial_i U(y)) \leqslant Ne^{\epsilon U(y)}$$

for an appropriate constant N which is independent of y, x. Besides ϕ and $\bar{\phi}$ are continuously differentiable on Q. If we use the notation $\partial_r^k f(x) := \partial_r^k f((x^{(1)},\cdots,x^{(M)})) := \frac{\partial f((x^{(1)},\cdots,x^{(M)}))}{\partial x_r^{(k)}}$ for $k = 1,\cdots,M$ and $r = 1,\cdots,d$, then for $x \in \mathbb{R}^{Md}$,

$$\bar{a}_{i,j}(t,x) = \sum_{k=1}^{M} a_{i-(k-1)d,j-(k-1)d}(x^{(k)}) I_{(k-1)d<i,j\leqslant kd}, \tag{5.62}$$

$$\partial_r^k \bar{a}_{i,j}(t,x) = \partial_r^k a_{i-(k-1)d,j-(k-1)d}(x^{(k)}) I_{(k-1)d<i,j\leqslant kd} = \partial_r a_{i-(k-1)d,j-(k-1)d}(x^{(k)}) I_{(k-1)d<i,j\leqslant kd}, \tag{5.63}$$

where $1 \leqslant i, j \leqslant Md$, and

$$\partial_r^k \bar{\phi}(t,x) = \sum_{1\leqslant q\neq k\leqslant M} \partial_r V((x^{(k)} - x^{(q)})sign(q-k))sign(q-k) + 4C \sum_{1\leqslant q\neq k\leqslant M} (x_r^{(k)} - x_r^{(q)}),$$

furthermore,

$$\begin{aligned}\partial_n^m \partial_r^k \bar{\phi}(t,x) = \sum_{1\leqslant q\neq k\leqslant M} \Big(& I_{m=k}\partial_n \partial_r V((x^{(k)} - x^{(q)})sign(q-k)) \\ & - I_{m=q}\partial_n\partial_r V((x^{(k)} - x^{(q)})sign(q-k))\Big) + 4C(I_{m=k,n=r} - I_{m\neq k,n=r}).\end{aligned}$$

Combining the above equalities with our assumptions of V and σ, by algebraic calculation we get that on Q there exists a large number $C_{M,d}$ depending on Md and a constant $C' \in (0,\infty)$ such that

$$\begin{aligned}& 2D_t\bar{\phi}(t,x) + \sum_{i,j=1}^{Md} \partial_j(\bar{a}_{i,j}\partial_i\bar{\phi})(t,x) \\ =& \sum_{i,j=1}^{d}\sum_{k=1}^{M} \Big(\partial_j^k a_{i,j}(x^{(k)})\partial_i^k\bar{\phi}(t,x) + a_{i,j}(x^{(k)})\partial_j^k\partial_i^k\bar{\phi}(t,x)\Big) \\ =& \sum_{i,j=1}^{d}\sum_{k=1}^{M}\sum_{1\leqslant q\neq k\leqslant M} \Big(\partial_j a_{i,j}(x^{(k)})[\partial_i V((x^{(k)} - x^{(q)})sign(q-k))sign(q-k) \\ & \qquad + 4C(x_i^{(k)} - x_i^{(q)})] \\ & \quad + a_{i,j}(x^{(k)})[\partial_j\partial_i V((x^{(k)} - x^{(q)})sign(q-k))]\Big) + \sum_{i,j=1}^{d}\sum_{k=1}^{M} a_{i,j}(x^{(k)})4CI_{i=j} \\ \leqslant & C_{M,d}\sum_{1\leqslant q<g\leqslant M}(Ce^{\epsilon(V(x^{(q)}-x^{(g)})+U(x^{(q)}-x^{(g)}))} + Ne^{\epsilon(U(x^{(q)}-x^{(g)}))}) \leqslant C'e^{\epsilon\bar{\phi}(t,x)}.\end{aligned}$$

The continuity of $\bar{a}_{i,j}(t,x)$ on Q and $\partial_j^k\bar{a}_{i,j}(t,x)$ on Q^n can be easily checked from (5.62) and (5.63) and conditions about σ. In order to reduce the lengthy algebraic computation, we only show the part for $\bar{a}_{i,j}(t,x)$, similarly we can get the desired continuity for $\partial_j^k\bar{a}_{i,j}(t,x)$ on Q^n. For any (t,x) and $(s,y) \in Q$, by (5.62) we have for $1 \leqslant i,j \leqslant Md$,

$$\begin{aligned}&|\bar{a}_{i,j}(t,x) - \bar{a}_{i,j}(s,y)| \\ &\qquad \leqslant C_{Md}\sum_{k=1}^{M}|a_{i-(k-1)d,j-(k-1)d}(x^{(k)}) - a_{i-(k-1)d,j-(k-1)d}(y^{(k)})|I_{(k-1)d<i,j\leqslant kd} \\ &\qquad \leqslant C_{Md}\sum_{k=1}^{M}|x^{(k)} - y^{(k)}| \\ &\qquad \leqslant C''|x-y|.\end{aligned}$$

We can adjust constants C'' and K such that there is still a positive constant such condition $(\sigma 1)$ satisfied.

It follows that all conditions on $\bar{\phi}$ and $\bar{\sigma}$ in Theorem 5.2 are fulfilled and therefore by Remark 5.14 the corresponding stochastic equation for a process $X_t = (X_t^{(1)}, ..., X_t^{(M)})$ has a unique strong solution defined for all times whenever for the initial condition x we have $(0, x) \in Q$. The corresponding equation is the following system

$$X_t^{(k)} = x^{(k)} + \int_0^t \sigma(X_s^{(k)}) dW_s^{(k)} - \int_0^t (\sigma\sigma^*)(X_s^{(k)}) \partial_k \bar{\phi}(s, X_s) ds + (\frac{1}{2} \sum_{j=1}^{d} \int_0^t \partial_j a_{i,j}(X_s^{(k)}) ds)_{1 \leqslant i \leqslant d} + \int_0^t b^{(k)}(s, X_s) ds.$$

We rewrite it as following with $k = 1, ..., M$

$$\begin{aligned} X_t^{(k)} = x^{(k)} &+ \int_0^t \sigma(X_s^{(k)}) dW_s^{(k)} \\ &- \int_0^t (\sigma\sigma^*)(X_s^{(k)}) \sum_{j=1, j\neq k}^{M} \nabla V((X_s^{(k)} - X_s^{(j)}) sign(j-k)) sign(j-k) ds \\ &+ (\frac{1}{2} \sum_{j=1}^{d} \int_0^t \partial_j a_{i,j}(X_s^{(k)}) ds)_{1 \leqslant j \leqslant d}, \end{aligned}$$

which has a unique strong solution defined for all times whenever $(0, (x^{(1)}, ..., x^{(M)})) \in Q$.

6 Existence and uniqueness of weak solutions to SDEs with distributional valued drifts and jump type noise

In this chapter, the well-posedness of nonlocal elliptic equation with singular drift is investigated in Besov-Hölder spaces. In the end, we show the existence and uniqueness for corresponding martingale problem, which is equivalent to show the existence and uniqueness for corresponding weak solution. Moreover, we prove that the one-dimensional distribution of the weak solution has a density in some Besov space.

6.1 Preliminaries and main results

We consider the following nonlocal elliptic equation in $\mathbb{R}^d$:

$$\lambda u - \mathscr{L}^\alpha_\kappa u - b \cdot \nabla u = f. \tag{6.1}$$

Here $\alpha \in (0,2)$, $b \in \mathscr{C}^\beta$(Besov-Hölder space, see Definition 6.7 below) with $\beta \in \mathbb{R}$, κ is a nonnegative measurable function from $\mathbb{R}^d \times \mathbb{R}^d$ to $[0,\infty)$ and

$$\mathscr{L}^\alpha_\kappa f(x) := \int_{\mathbb{R}^d} \left(f(x+z) - f(x) - \nabla f(x) \cdot z^{(\alpha)}\right) \frac{\kappa(x,z)}{|z|^{d+\alpha}} \mathrm{d}z,$$

where $z^{(\alpha)} := z\mathbf{1}_{\{|z|<1\}}\mathbf{1}_{\alpha=1} + z\mathbf{1}_{\alpha\in(1,2)}$.

The first aim of our work is to establish a Schauder's type estimate for the solution to (6.1) with irregular coefficients. There are many literatures studied this problem in different settings. When $\alpha \in (1,2)$, b is a Hölder continuous function and $\mathscr{L}^\alpha_\kappa$ is some α-stable type operator, Priola in [52] and [53] studied the a priori estimate by using classic perturbation argument. Similarly, Athreya, Butkovsky and Mytnik in [5] showed the global estimate for $\mathscr{L}^\alpha_\kappa = \Delta^{\alpha/2}$ with $\alpha \in (1,2)$ and $b \in \mathscr{C}^\beta$ with $\beta > \frac{1-\alpha}{2}$. Indeed, the analytic result in [5] also holds for any non degenerate α-stable operators. For $\alpha > 1$, in [48], Mikulevicius and Pragarauskas also studied the nonlocal Cauchy problem with first order term in Hölder space. And recently, in [17], Dong, Jin and Zhang studied the Dini and Schauder estimate for nonlocal fully nonlinear equations. However, when $\alpha < 1$, both [48] and [17] must assume $b \equiv 0$. To our best knowledge, when $\alpha \in (0,1)$, the interior estimate for the solution to (6.1) with non divergence free drift was first obtained by Silvestre in [57]. He used the extension method for $\mathscr{L}^\alpha_\kappa = \Delta^{\alpha/2}$ when $\alpha \in (0,1)$ and $b \in \mathscr{C}^\beta$ with $\beta > 1-\alpha$ to reduce the nonlocal problem to the local case. Recently, similar result was extended for stable-like operators in [78] by using Littlewood-Paley theory. Let us also mention that there are much more works for nonlocal equation without first order term, for instance [6], [16] and the references therein.

In this work, we will show the global estimates in more general setting. Our assumption on κ is:

Assumption 5. *There are constants $r_0, \Lambda_1, \Lambda_2, \Lambda_3 > 0$, $\vartheta \in (0,1)$ such that*

$$\int_{B_r} \kappa(x,z)\,\mathrm{d}z \geqslant \Lambda_1 r^d, \quad x \in \mathbb{R}^d, r \in (0,r_0]; \tag{$\mathbf{H}_1$}$$

$$\kappa(x,z) \leqslant \Lambda_2,\ x,z \in \mathbb{R}^d;\quad \mathbf{1}_{\alpha=1} \int_{\{r<|z|<R\}} z \cdot \kappa(x,z) \mathrm{d}z = 0,\ 0<r<R<\infty; \qquad (\mathbf{H}_2)$$

$$|\kappa(x,z) - \kappa(y,z)| \leqslant \Lambda_3 |x-y|^\vartheta, \quad x,y \in \mathbb{R}^d, \vartheta \in (0,1). \qquad (\mathbf{H}_3)$$

The following is our first main result:

Theorem 6.1. *Suppose $\kappa(x,z)$ satisfies ($\mathbf{H}_1$)-($\mathbf{H}_3$) and* $\max\{0,(1-\alpha)\} < \vartheta < 1$.

1. *If $\alpha \in (0,1]$, $\beta \in (1-\alpha,\vartheta)$ and $b \in \mathscr{C}^\beta$, then there are constants $\lambda_0, C > 0$ such that for any $\lambda \geqslant \lambda_0$ and $f \in \mathscr{C}^\beta$, equation (6.1) has a unique solution in $\mathscr{C}^{\alpha+\beta}$ satisfying*

$$(\lambda - \lambda_0)\|u\|_{\mathscr{C}^\beta} + \|u\|_{\mathscr{C}^{\alpha+\beta}} \leqslant C\|f\|_{\mathscr{C}^\beta}, \qquad (6.2)$$

where λ_0, C only depend on $d,\alpha,\beta,\vartheta,r_0,\Lambda_1,\Lambda_2,\Lambda_3,\|b\|_{\mathscr{C}^\beta}$.

2. *If $\alpha \in (1,2)$, $\beta \in (-(\frac{\alpha-1}{2} \wedge \vartheta),\vartheta)$ and $b \in \mathscr{C}^\beta$, then the above conclusions also hold.*

Notice that our condition ($\mathbf{H}_1$) is much weaker than the usual lower bounded assumption $\kappa(x,z) \geqslant \lambda > 0$ and also weaker than Assumption A(i) in [48]. A typical example is take

$$\kappa(x,z) = \mathbf{1}_{V(x)}(z).$$

Here $V(x) \in \mathbb{R}^d$ is a conical set of the form $V(x) = \{z \in \mathbb{R}^d : |\langle z/|z|, \xi(x)\rangle| > \delta\}$ with measurable $\xi : \mathbb{R}^d \to \mathbb{S}^{d-1}$, and $\delta > 0$ is fixed.

Like in [78], our approach of getting the Schauder type estimate is based on Littlewood-Paley theory. For the first case in Theorem 6.1, the key step is to establish a frequency localized maximum inequality(see Lemma 6.11 below). This kind of maximum principle appeared in [66] for $\kappa \equiv 1$. We extend their result for any $\kappa(x,z) = \kappa(z)$ satisfying ($\mathbf{H}_4$) below. When $\alpha > 1$ and $\beta \in (-(\frac{\alpha-1}{2} \wedge \vartheta), 0]$, the main problem is how to prove the boundedness of $\mathscr{L}^\alpha_\kappa : \mathscr{C}^{\alpha+\beta} \to \mathscr{C}^\beta$, where the Bony's decomposition plays a crucial rule in our proof.

As one of the motivations of considering the regularity estimate for (6.1), we want to investigate the well-posedness of the following SDE in $\mathbb{R}^d$:

$$X_t = X_0 + \int_0^t \sigma(X_{s-})\mathrm{d}Z_s + \int_0^t b(X_s)\mathrm{d}s \qquad (6.3)$$

in weak sense. Here Z_t is an α-stable process in $\mathbb{R}^d$, σ is a $d \times d$-matrix-valued measurable function and b is the drift, which might be very singular. Suppose Z_t is rotational symmetric, $L^\alpha_\sigma + b\cdot\nabla$ is the generator of X_t, for any σ satisfies (6.5) below, we have

$$\begin{aligned} L^\alpha_\sigma f(x) + b\cdot\nabla f(x) &= \int_{\mathbb{R}^d} (f(x+\sigma(x)z) - f(x) - \nabla f(x)\cdot\sigma(x)z^{(\alpha)}) \frac{\mathrm{d}z}{|z|^{d+\alpha}} + b\cdot\nabla f(x) \\ &= \int_{\mathbb{R}^d} (f(x+z) + f(x) - \nabla f(x)\cdot z^{(\alpha)}) \frac{\mathrm{d}z}{|\det\sigma(x)|\cdot|\sigma^{-1}(x)z|^{d+\alpha}} \\ &\quad + b\cdot\nabla f(x) = \mathscr{L}^\alpha_{\kappa,b} f(x), \end{aligned}$$

where

$$\kappa(x,z) := \frac{|z|^{d+\alpha}}{|\det\sigma(x)|\cdot|\sigma^{-1}(x)z|^{d+\alpha}}. \tag{6.4}$$

Since the well-posedness of the resolvent equations or backward Kolmogorov equations associated with $L^\alpha_\sigma + b\cdot\nabla$ are closely related to the weak solutions(or martingale solutions) of (6.3), our analytic result Theorem 6.1 has direct applications to SDE driven by α-stable process.

On the other hand, pathwise uniqueness and strong existence for (6.3) with irregular coefficients have already been studied in a large number of literatures, see [60] for one dimensional case and [52], [73], [53], [12], [14], etc for more general Lévy noises in $\mathbb{R}^d$. Roughly speaking, these works showed that the SDE (6.3) has a unique strong solution under the conditions that σ is bounded, uniformly nondegenerate and Lipschitz, Z_t is a non degenerated α-stable process, $b \in \mathscr{C}^\beta$ with $\beta > 1-\frac{\alpha}{2}$. However, when we consider the existence and uniqueness of weak solutions to (6.3) or the well-posedness of corresponding martingale problem, the regularity assumptions on the coefficients can be released. In [76], the authors considered (6.3) driven by Brownian motion, they showed that if $\sigma = \mathbb{I}$, $b \in H_p^{-\frac{1}{2}}$ with $p > 2d$ one can still give a natural meaning of "$\int_0^t b(X_s)\mathrm{d}s$"(see also [77]). The drift term may not be a process with finite variation any more but an additive functional of X with zero energy. In [5], they considered the similar SDEs driven by one dimensional additive α-stable noise with singular drifts in Besov-Hölder space. The above works are motivated by Bass and Chen's earlier works [9], [10].

In this chapter, we will study the martingale problem associated with $\mathscr{L}^\alpha_{\kappa,b} := \mathscr{L}^\alpha_\kappa + b\cdot\nabla$. When $\alpha \leqslant 1$, since we assume $b \in \mathscr{C}^\beta$ with $\beta > 0$, there is no issue about the definition of martingale or weak solution. However, when $\alpha > 1$ and $b \in \mathscr{C}^\beta$ with $\beta \leqslant 0$, like in [76], [5], we need to give an appropriate definition of solutions to (6.3)(see Definition 6.22). Combining Theorem 6.1 and some standard techniques in probability theory, we can obtain the following result. We distribute the proof in Lemma 6.23 and Lemma 6.24.

Corollary 6.2. *Suppose* $\max\{0,(1-\alpha)\} < \vartheta < 1$, $\kappa(x,z)$ *satisfies* (**H**$_1$)-(**H**$_3$), *and* $b \in \mathscr{C}^\beta$, *where* $\beta \in (1-\alpha,\vartheta)$ *if* $\alpha \in (0,1]$ *and* $\beta \in (-(\frac{\alpha-1}{2}\wedge\vartheta),0]$ *if* $\alpha \in (1,2)$. *Then, for each* $x \in \mathbb{R}^d$, *there is a unique probability measure* $\mathbb{P}_x$ *with starting point* x *on the Skorokhod space* $\mathbb{D}$, *which solves the martingale problem associated with* $\mathscr{L}^\alpha_{\kappa,b}$ *and satisfies the Krylov's type estimate(see Definition 6.20).*

Our corollary above implies:

Proposition 6.3. *Suppose* Z_t *is a rotational symmetric* α*-stable process,* σ *satisfies*

$$\Lambda^{-1}|z| \leqslant |\sigma(x)z| \leqslant \Lambda|z|, \quad \Lambda > 0, z \in \mathbb{R}^d. \tag{6.5}$$

then

(i) If $\alpha \in (0,1]$, $\beta \in (1-\alpha,1)$. $\sigma, b \in \mathscr{C}^\beta$, *there is a unique weak solution to* (6.3).

(ii) *If* $\alpha \in (1,2)$, $\beta \in (\frac{1-\alpha}{2}, 0]$, $\varepsilon > 0$. $\sigma \in \mathscr{C}^{-\beta+\varepsilon}$, $b \in \mathscr{C}^{\beta}$, *there is a unique weak solution to* (6.3).

Another interesting problem we attempt to study in this chapter is the regularity estimates for the one dimensional distribution of the solutions to martingale problem associated with $\mathscr{L}^{\alpha}_{\kappa,b}$. Debussche and Fournier in [15] proved that the law of the solution to (6.3) has a density in some Besov space, under some non-degeneracy condition on the driving Lévy process and some Hölder-continuity assumptions on the coefficients. We following the arguments in [15], but instead of using the crucial Lemma 2.1 therein, we use the Littlewood-Paley description of Besov spaces to simplify the proof and get a bit more general result(see Lemma 6.30).

Theorem 6.4. *Under the same conditions in Corollary 6.2 for each* $x \in \mathbb{R}^d$, *suppose* $\mathbb{P}_x$ *is the unique solution in Corollary 6.2. Then, for each* $t > 0$ *the distribution of canonical process* ω_t *under* $\mathbb{P}_x$ *has a density in Besov space* $B^{\gamma}_{q,\infty}$ *with* γ *and* q *satisfying*

$$0 < \gamma < \alpha(\alpha+\beta-1), \quad 1 \leqslant q < \frac{d}{d+\gamma-\alpha(\alpha+\beta-1)}$$

if $\alpha \in (0,1]$ *and*

$$0 < \gamma < (\alpha+\beta-1) \wedge \tfrac{\vartheta}{\alpha}, \quad 1 \leqslant q < \frac{d}{d+\gamma-(\alpha+\beta-1)\wedge\frac{\vartheta}{\alpha}}$$

if $\alpha \in (1,2)$.

Let $\mathscr{P}(\mathbb{R}^d)$ be the collection of all probability measures on $\mathbb{R}^d$. Combining Corollary 6.2 and Theorem 6.4, we obtain the following interesting corollary:

Corollary 6.5. *Under the same conditions in Corollary 6.2 for any* $x \in \mathbb{R}^d$, *the following nonlocal Fokker-Planck equation:*

$$\langle \varrho_t, \phi\rangle = \phi(x) + \int_0^t \langle \varrho_s, \mathscr{L}^{\alpha}_{\kappa,b}\phi\rangle \mathrm{d}s, \quad \forall \phi \in C_c^\infty \tag{6.6}$$

has a unique *solution* $\{\varrho_t\} \subseteq \mathscr{P}(\mathbb{R}^d)$. *Moreover, for each* $t > 0$, $\varrho_t \in B^{\gamma}_{q,\infty}$ *with* γ *and* q *satisfying* $0 < \gamma < \alpha(\alpha+\beta-1)$, $1 \leqslant q < \frac{d}{d+\gamma-\alpha(\alpha+\beta-1)}$ *if* $\alpha \in (0,1]$ *and* $0 < \gamma < (\alpha+\beta-1)\wedge\frac{\vartheta}{\alpha}$, $1 \leqslant q < \frac{d}{d+\gamma-(\alpha+\beta-1)\wedge\frac{\vartheta}{\alpha}}$ *if* $\alpha \in (1,2)$.

Remark 6.6. *The above result can also be seen as a probabilistic approach to the theory of regularity of solutions to non-local partial differential equations. We give a probabilistic proof for the well-posedness as well as regularity estimates for linear Fokker-Plank equation with singular coefficients and initial data.*

This chapter is organized as follows: In Section 6.2, we recall some basic knowledge from Littlewood-Paley theory for later use. We establish apriori estimates for (6.1) in Hölder-Besov spaces in Section 6.3. In Section 6.4, we prove the well-posedness of martingale problem associated with $\mathscr{L}^{\alpha}_{\kappa,b}$. In Section 6.5, we show the one dimensional distribution of the martingale solution has a density in some Besov space. For the completeness of the paper we add the equivalence between martingale solution and weak solution in Appendix.

6.2 Preparations

In this section, we recall some basic concepts and properties of Littlewood-Paley decomposition that will be used later.

Let $\mathscr{S}(\mathbb{R}^d)$ be the Schwartz space of all rapidly decreasing functions, and $\mathscr{S}'(\mathbb{R}^d)$ the dual space of $\mathscr{S}(\mathbb{R}^d)$. Given $f \in \mathscr{S}(\mathbb{R}^d)$, let $\mathscr{F} f = \hat{f}$ be the Fourier transform of f defined by

$$\hat{f}(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x) \mathrm{d}x.$$

Let $\chi : \mathbb{R}^d \to [0, 1]$ be a smooth radial function with

$$\chi(\xi) = 1, \ |\xi| \leqslant 1, \ \chi(\xi) = 0, \ |\xi| \geqslant 3/2.$$

Define

$$\mathcal{C} := B_{3/2} \backslash \overline{B_{1/2}} = \{x \in \mathbb{R}^d : 1/2 < |x| < 3/2\}; \quad \varphi(\xi) := \chi(\xi) - \chi(2\xi).$$

It is easy to see that $\varphi \geqslant 0$ and supp $\varphi \subset \mathcal{C}$ and

$$\chi(2\xi) + \sum_{j=0}^{k} \varphi(2^{-j}\xi) = \chi(2^{-k}\xi) \overset{k\to\infty}{\to} 1. \tag{6.7}$$

In particular, if $|j - j'| \geqslant 2$, then

$$\mathrm{supp}\varphi(2^{-j}\cdot) \cap \mathrm{supp}\varphi(2^{-j'}\cdot) = \emptyset.$$

In this paper we shall fix such χ and φ and also introduce another nonnegative function $\tilde{\varphi} \in C_c^\infty(\mathbb{R}^d)$ supported on $B_2 \backslash \overline{B_{1/4}}$ and $\tilde{\varphi} = 1$ on $\mathcal{C}$ for later use.

We introduce the definition of Besov space below.

Definition 6.7. *The dyadic block operator Δ_j is defined by*

$$\Delta_j f := \begin{cases} \mathscr{F}^{-1}(\chi(2\cdot)\mathscr{F} f), & j = -1, \\ \mathscr{F}^{-1}(\varphi(2^{-j}\cdot)\mathscr{F} f), & j \geqslant 0. \end{cases}$$

For $s \in \mathbb{R}$ and $p, q \in [1, \infty]$, the Besov space $B^s_{p,q}$ is defined as the set of all $f \in \mathscr{S}'(\mathbb{R}^d)$ with

$$\|f\|_{B^s_{p,q}} := \left(\sum_{j \geqslant -1} 2^{jsq} \|\Delta_j f\|_p^q \right)^{1/q} < \infty;$$

If $p = q = \infty$, we denote $\mathscr{C}^s := B^s_{\infty,\infty}$.

Let

$$h := \mathscr{F}^{-1}\varphi, \quad \tilde{h} := \mathscr{F}^{-1}\tilde{\varphi}, \quad h_{-1} := \mathscr{F}^{-1}\chi(2\cdot);$$

$$h_j := \mathscr{F}^{-1}\varphi(2^{-j}\cdot) = 2^{jd} h(2^j\cdot), \quad j \geqslant 0.$$

By definition it is easy to see that

$$\Delta_j f(x) = (h_j * f)(x) = \int_{\mathbb{R}^d} h_j(x - y) f(y) \mathrm{d}y, \quad j \geqslant -1. \tag{6.8}$$

Definition 6.8. *The low-frequency cut-off operator S_j is defined by*

$$S_j f := \sum_{j' \leqslant j-1} \Delta_{j'} f.$$

The paraproduct of f and g is defined by

$$T_f g := \sum_{j \geqslant -1} S_{j-1} f \Delta_j g.$$

The remainder of f and g is defined by

$$R(f,g) = \sum_{|k-j| \leqslant 1} \Delta_k f \Delta_j g.$$

The following two Lemmas can be found in [61].

Lemma 6.9. *If $s > 0, s \notin \mathbb{N}$, then*

$$\mathscr{C}^s := B^s_{\infty,\infty} \asymp C^s,$$

where C^s is the usual Hölder space.

Lemma 6.10 (Bernstein's inequalities)**.** *For any $1 \leqslant p \leqslant q \leqslant \infty$ and $j \geqslant 0$, we have*

$$\|\nabla^k \Delta_j f\|_q \leqslant C_p 2^{(k+d(\frac{1}{p}-\frac{1}{q}))j} \|\Delta_j f\|_p, \ k = 0, 1, \cdots, \tag{6.9}$$

and

$$\|(-\Delta)^{s/2} \Delta_j f\|_q \leqslant C_p 2^{(s+d(\frac{1}{p}-\frac{1}{q}))j} \|\Delta_j f\|_p, \quad s \geqslant 0. \tag{6.10}$$

6.3 Schauder estimates for (6.1)

In this section, we establish the Schauder type estimate for (6.1) and its well-posedness in Besov-Hölder space.

6.3.1 The case $\kappa(x,z) = \kappa(z)$

The following assumptions will be needed in this subsection.

Assumption 6. *There are constants $r_0, \delta_0, \Lambda, \Lambda_2 > 0$ such that*

$$|\{z : \kappa(z) > \Lambda\} \cap B_r| \geqslant \delta_0 r^d, \quad r \in (0, r_0]; \tag{$\mathbf{H}_4$}$$

$$\kappa(z) \leqslant \Lambda_2, \ z \in \mathbb{R}^d; \ \ \mathbf{1}_{\alpha=1} \int_{\{r<|z|<R\}} z \cdot \kappa(z) \mathrm{d}z = 0, \ 0 < r < R < \infty. \tag{$\mathbf{H}_5$}$$

Recalling that $\mathcal{C} := \{x \in \mathbb{R}^d : 1/2 < |x| < 3/2\}$. Define

$$\mathscr{B} = \Big\{u \in \mathscr{S}(\mathbb{R}^d) : \text{supp } \hat{u} \in \mathcal{C}\Big\}, \quad J(u) = \Big\{x \in \mathbb{R}^d : |u(x)| = \|u\|_\infty\Big\}.$$

We have the following important frequency localized maximum principle.

Lemma 6.11. *There exists a number $c = c(d, \alpha, r_0, \delta_0) > 0$ such that for any κ satisfying ($\mathbf{H}_4$), the following maximal inequality holds:*

$$\inf_{u\in\mathscr{B}} \inf_{x\in J(u)} \Big\{\text{sgn}(u(x)) \cdot (-\mathscr{L}^\alpha_\kappa u(x))\Big\} \geqslant c \cdot \Lambda \|u\|_\infty. \tag{6.11}$$

The following simple lemma is needed in the proof of Lemma 6.11.

Lemma 6.12. *Suppose f is a real analytic function on $\mathbb{R}^d$, if f vanishes on a measurable subset of $\mathbb{R}^d$ whose Lebesgue measure is positive, then $f \equiv 0$ on $\mathbb{R}^d$.*

Proof. We prove the lemma by induction. Let m_k be the Lebesgue measure on $\mathbb{R}^k$.

- If $d = 1$, then f is analytic with $f|_E \equiv 0$ and $m_1(E) > 0$, which implies zero points of f must have an accumulation point on the line, by identity theorem, $f \equiv 0$.
- Assume the claim holds for $d - 1$. If $m_d(E) > 0$, then by Fubini theorem, there is a set $E_1 \subset \mathbb{R}$ with $m_1(E_1) > 0$, such that for any $x_1 \in E_1$,

  $$m_{d-1}(E \cap \{x_1\} \times \mathbb{R}^{d-1}) > 0.$$

  By induction hypothesis, for each $x_1 \in E_1$, function $z \mapsto f(x_1, z)$ vanishes identically. Since $m_1(E_1) > 0$, we can find $x_1^n \in E_1$, $x_1^n \to a$. Now for each $z \in \mathbb{R}^{d-1}$, function $x_1 \mapsto f(x_1, z)$ is real analytic, its zero points has an accumulation point a. By the conclusion for 1 dimensional case, we get $f(x_1, z) \equiv 0$.

□

Proof of Lemma 6.11. Without loss of generality, we can assume $\Lambda = 1$. Define

$$\mathscr{A}(r_0, \delta_0) := \{\kappa : \kappa \text{ satisfies } (\mathbf{H}_4) \text{ with } \Lambda = 1\},$$

$$c := \inf_{\kappa\in\mathscr{A}(r_0,\delta_0)} \inf_{u\in\mathscr{B}} \inf_{x\in J(u)} \Big\{\text{sgn}(u(x)) \cdot (-\mathscr{L}^\alpha_\kappa u(x))\Big\}/\|u\|_\infty.$$

We emphasize that the constant c only depends on d, α, r_0, δ_0. By the definition of c, there exists a sequence of smooth functions $w_n \in \mathscr{S}(\mathbb{R}^d)$ satisfying $\text{supp}\, \hat{w}_n \subset \mathcal{C}$, $x_n \in J(w_n)$ and $\kappa_n(z) \in \mathscr{A}(r_0, \delta_0)$ such that

$$w_n(x_n) = \max_{x\in\mathbb{R}^d} |w_n| = 1, \quad \lim_{n\to\infty} \big[-\mathscr{L}^\alpha_{\kappa_n} w_n(x_n)\big] = c.$$

Let $u_n(x) := w_n(x_n + x)$, it's easy to see that $u_n \in \mathscr{B}$ and

$$u_n(0) = \max_{x\in\mathbb{R}^d} |u_n|(x) = 1, \quad \lim_{n\to\infty} \big[-\mathscr{L}^\alpha_{\kappa_n} u_n(0)\big] = c. \tag{6.12}$$

Notice that

$$u_n(x) = \int_{\mathbb{R}^d} \tilde{h}(x-y)u_n(y)\mathrm{d}y,$$

where $\tilde{h}$ is defined in section 2. For any $k \in \mathbb{N}$,

$$\|\nabla^k u_n\|_\infty = \|\nabla^k \tilde{h} * u_n\|_\infty \leqslant \|\nabla^k \tilde{h}\|_1 \|u_n\|_\infty \leqslant C_k.$$

By Ascoli-Azela's lemma and diagonal argument, there is a subsequence of $\{u_n\}$(still denoted by u_n for simple) and $u \in C_b^\infty$ such that $\nabla^k u_n$ converges to $\nabla^k u$ uniformly on any compact set. Let $\chi_R(\cdot) = \chi(\cdot/R)$, where χ is the same function in section 2. For any $\phi \in \mathscr{S}(\mathbb{R}^d)$,

$$\begin{aligned}\left|\int \phi(u_n - u)\right| \leqslant& \int |\phi\chi_R \cdot (u_n-u)| + \int |\phi(1-\chi_R)(u_n-u)| \\ \leqslant& \|\phi\|_{L^1}\|u_n - u\|_{L^\infty(B_{3R/2})} + 2\sup_{|x|>R}|\phi(x)|.\end{aligned}$$

Let $n \to \infty$ and then $R \to \infty$, we get

$$\langle \phi, u_n\rangle \to \langle \phi, u\rangle, \quad \forall \phi \in \mathscr{S}(\mathbb{R}^d).$$

i.e. $u_n \to u$ in $\mathscr{S}'(\mathbb{R}^d)$ and consequently, $\hat{u}_n \to \hat{u}$ in $\mathscr{S}'(\mathbb{R}^d)$. For any $\phi \in \mathscr{S}(\mathbb{R}^d)$ supported on $\mathbb{R}^d \backslash \mathcal{C}$, we have

$$\langle \phi, \hat{u}\rangle = \lim_{n\to\infty} \langle \phi, \hat{u}_n\rangle = 0,$$

which means u is also supported on $\mathcal{C}$. Thus the complex-valued function

$$U : z \mapsto (2\pi)^{-d}\langle \mathrm{e}^{i\langle z,\xi\rangle}, \hat{u}\rangle$$

is a holomorphic function on $\mathbb{C}^d$ and $u = U|_{\mathbb{R}^d}$. This implies u is a real analytic function.

Now assume $c = 0$, for any $\lambda \in (0,1)$, by (6.12) and the fact that $\nabla u_n(0) = 0$, we have

$$\begin{aligned}-\mathscr{L}^\alpha_{\kappa_n} u_n(0) =& \int_{B_{r_0}} (u_n(0) - u_n(z))\frac{\kappa_n(z)}{|z|^{d+\alpha}}\mathrm{d}z \\ \geqslant& \int_{B_{r_0}\cap\{u_n\leqslant\lambda\}\cap\{\kappa_n>1\}} (1-u_n(z))\frac{\mathrm{d}z}{|r_0|^{d+\alpha}} \\ \geqslant& (1-\lambda)r_0^{-d-\alpha}|B_{r_0}\cap\{u_n \leqslant \lambda\}\cap\{\kappa_n > 1\}|.\end{aligned}$$

This yields

$$\limsup_{n\to\infty}|B_{r_0}\cap\{u_n\leqslant\lambda\}\cap\{\kappa_n>1\}| \leqslant (1-\lambda)^{-1}r_0^{d+\alpha}\lim_{n\to\infty}[-\mathscr{L}^\alpha_{\kappa_n}u_n(0)] = 0.$$

Combining the above estimate and our assumption (**H**$_4$), we get

$$\begin{aligned}&\liminf_{n\to\infty}|B_{r_0}\cap\{u_n>\lambda\}\cap\{\kappa_n>1\}| \\ =&\liminf_{n\to\infty}|B_{r_0}\cap\{\kappa_n>1\}| - \limsup_{n\to\infty}|B_{r_0}\cap\{u_n\leqslant\lambda\}\cap\{\kappa_n>1\}| \geqslant \delta_0 r_0^d.\end{aligned}$$

One the other hand, $u_n \to u$ uniformly in B_{r_0} implies

$$|B_{r_0} \cap \{u > \lambda\}| = \lim_{n\to\infty} |B_{r_0} \cap \{u_n > \lambda\}| \\ \geqslant \liminf_{n\to\infty} |B_{r_0} \cap \{u_n > \lambda\} \cap \{\kappa_n > 1\}| \geqslant \delta_0 r_0^d.$$

Notice that $u \leqslant 1$, let $\lambda \uparrow 1$ in the first term above, we obtain $|\{x \in B_{r_0} : u(x) = 1\}| \geqslant \delta_0 r_0^d > 0$. Using Lemma 6.12, we obtain $u \equiv 1$ on $\mathbb{R}^d$ i.e. $\hat{u} = \delta_0$, the Dirac measure. However, as we see before, $\hat{u}$ must be supported on $\mathcal{C}$, this contradiction implies $c = c(d, \alpha, r_0, \delta_0) > 0$. □

Corollary 6.13. *Let $R \geqslant 1$. Suppose κ satisfies ($\mathbf{H}_4$) and* supp $\hat{u} \subset R\mathcal{C} := \{x : x/R \in \mathcal{C}\}$, *then there is a positive constant $c = c(d, \alpha, r_0, \delta_0)$ such that*

$$\inf_{x\in J(u)} \Big\{\mathrm{sgn}(u(x)) \cdot (-\mathscr{L}^\alpha_\kappa u(x))\Big\} \geqslant c\,\Lambda R^\alpha \|u\|_\infty, \tag{6.13}$$

where $J(u) = \{x : |u(x)| = \|u\|_\infty\}$.

Proof. Suppose $x_0 \in J(u)$, define $u^{x_0}_R(x) := u(x_0 + x/R)$, $\kappa_R(z) := R^\alpha \kappa(z/R)$. By our assumption on u, one can see that supp $\widehat{u^{x_0}_R} \subset \mathcal{C}$ and κ_R satisfies ($\mathbf{H}_4$) with constant Λ replaced by ΛR^α. Notice that

$$\mathscr{L}^\alpha_{\kappa_R} u^{x_0}_R(0) = \int_{\mathbb{R}^d} (u(x_0 + z/R) - u(x_0)) \frac{\kappa(z/R)}{|z/R|^{d+\alpha}} \mathrm{d}(z/R) \\ = \int_{\mathbb{R}^d} (u(x_0 + z) - u(x_0)) \frac{\kappa(z)}{|z|^{d+\alpha}} \mathrm{d}z = \mathscr{L}^\alpha_\kappa u(x_0),$$

by Lemma 6.11, we obtain that

$$-\mathrm{sgn}(u(x_0)) \cdot \mathscr{L}^\alpha_\kappa u(x_0) = -\,\mathrm{sgn}(u^{x_0}_R(0)) \cdot \mathscr{L}^\alpha_{\kappa_R} u^{x_0}_R(0) \\ \geqslant c\,\Lambda R^\alpha \|u^{x_0}_R\|_\infty = c\,\Lambda R^\alpha \|u\|_\infty.$$

So we complete our proof. □

We need the following simple commutator estimate.

Lemma 6.14. *For any $j \geqslant -1$, $\beta \in (0,1)$,*

$$\|[\Delta_j, b \cdot \nabla]u\|_\infty \leqslant C 2^{-\beta j} \|b\|_{\mathscr{C}^\beta} \|\nabla u\|_{L^\infty}, \tag{6.14}$$

where $C = C(d, \beta)$.

Proof. By (6.8), we have

$$[\Delta_j, b \cdot \nabla]u(x) = \int_{\mathbb{R}^d} h_j(y)(b(x-y) - b(x)) \cdot \nabla u(x-y) \mathrm{d}y,$$

hence for any $\beta \in (0,1)$,

$$
\begin{aligned}
\|[\Delta_j, b\cdot\nabla]u\|_\infty &\leqslant \int_{\mathbb{R}^d} h_j(y)\|b(\cdot-y)-b(\cdot)\|_\infty\|\nabla u\|_{L^\infty}\mathrm{d}y \\
&\leqslant C\|b\|_{\mathscr{C}^\beta}\|\nabla u\|_{L^\infty}\int_{\mathbb{R}^d}|h_j(y)|\,|y|^\beta\mathrm{d}y \\
&= C\|b\|_{\mathscr{C}^\beta}\|\nabla u\|_{L^\infty}2^{-\beta j}\int_{\mathbb{R}^d}|2h(2y)-h(y)|\,|y|^\beta\mathrm{d}y \\
&\leqslant C2^{-\beta j}\|b\|_{\mathscr{C}^\beta}\|\nabla u\|_{L^\infty}.
\end{aligned} \tag{6.15}
$$

□

Theorem 6.15. *Suppose κ satisfies ($\mathbf{H}_4$), if $\alpha \in (0,1]$ and $b \in \mathscr{C}^\beta$ with $\beta \in (1-\alpha, 1)$; or $\alpha \in (1,2)$ and $b \in \mathscr{C}^\beta$ with $\beta \in (-\frac{\alpha-1}{2}, 1)$. Then there a constant λ_0 such that for any $\lambda \geqslant \lambda_0$ and $f \in \mathscr{C}^\beta$ (6.1) has a unique solution in $\mathscr{C}^{\alpha+\beta}$. Moreover, we have the following apriori estimate*

$$
(\lambda-\lambda_0)\|u\|_{\mathscr{C}^\beta} + \|u\|_{\mathscr{C}^{\alpha+\beta}} \leqslant C\|f\|_{\mathscr{C}^\beta}, \tag{6.16}
$$

here $C = C(d,\alpha,\beta,r_0,\delta_0,\Lambda,\|b\|_{\mathscr{C}^\beta}) > 0$, $\lambda_0 = \lambda_0(d,\alpha,\beta,r_0,\delta_0,\Lambda,\|b\|_{\mathscr{C}^\beta}) \geqslant 0$.

Proof. For $\alpha \in (0,1]$, we frist assume $u \in \mathscr{S}(\mathbb{R}^d)$. Notice that $\Delta_j\mathscr{L}^\alpha_\kappa = \mathscr{L}^\alpha_\kappa\Delta_j$, we have

$$
\lambda\Delta_j u - \mathscr{L}^\alpha_\kappa\Delta_j u - b\cdot\nabla\Delta_j u = \Delta_j f + [\Delta_j, b\cdot\nabla]u.
$$

For $j=-1$, then

$$
\lambda\Delta_{-1}u - \mathscr{L}^\alpha_\kappa\Delta_{-1}u - b\cdot\nabla\Delta_{-1}u = \Delta_{-1}f + [\Delta_{-1}, b\cdot\nabla]u.
$$

Suppose $\Delta_{-1}u(x_{-1}) = \|\Delta_{-1}u\|_\infty$, noticing $\mathscr{L}^\alpha_\kappa\Delta_{-1}u(x_{-1}) \leqslant 0$ and $\nabla u(x_{-1}) = 0$, we get

$$
\begin{aligned}
\lambda\|\Delta_{-1}u\|_\infty \leqslant&\lambda\Delta_{-1}u(x_{-1}) - \mathscr{L}^\alpha_\kappa\Delta_{-1}u(x_{-1}) \\
\leqslant&\|\Delta_{-1}f\|_\infty + \|[\Delta_{-1}, b\cdot\nabla]u\|_\infty \\
\leqslant&\|\Delta_{-1}f\|_\infty + C\|b\|_{\mathscr{C}^\beta}\|u\|_{B^1_{\infty,1}}.
\end{aligned}
$$

For $j \geqslant 0$, assume $\mathrm{sgn}(\Delta_j u(x_j))\cdot\Delta_j u(x_j) = \|\Delta_j u\|_\infty$, by Lemma 6.11

$$
\begin{aligned}
(\lambda + c2^{\alpha j}))\|\Delta_j u\|_\infty =&\mathrm{sgn}(\Delta_j u(x_j))\cdot[\lambda\Delta_j u(x_j) + c2^{\alpha j}\Delta_j u(x_j)] \\
\leqslant&\|\lambda\Delta_j u - \mathscr{L}^\alpha_\kappa\Delta_j u - b\cdot\nabla\Delta_j u\|_\infty \\
\leqslant&\|\Delta_j f\|_\infty + \|[\Delta_j, b\cdot\nabla]u\|_\infty \\
\leqslant&\|\Delta_j f\|_\infty + C2^{-\beta j}\|b\|_{\mathscr{C}^\beta}\|u\|_{B^1_{\infty,1}}.
\end{aligned}
$$

Combining the above inequalities and using interpolation,

$$
\big(\lambda 2^{\beta j} + c2^{(\alpha+\beta)j}\big)\|\Delta_j u\|_\infty \leqslant 2^{\beta j}\|\Delta_j f\|_\infty + \|b\|_{\mathscr{C}^\beta}\big(\varepsilon\|u\|_{\mathscr{C}^{\alpha+\beta}} + C_\varepsilon\|u\|_{\mathscr{C}^\beta}\big),
$$

hence,

$$(\lambda - C_\varepsilon)\|u\|_{\mathscr{C}^\beta} + (c - \varepsilon C\|b\|_{\mathscr{C}^\beta})\|u\|_{\mathscr{C}^{\alpha+\beta}} \leqslant \|f\|_{\mathscr{C}^\beta}.$$

Choosing ε_0 sufficiently small, such that $(c - \varepsilon_0\|b\|_{\mathscr{C}^\beta}) \geqslant \frac{c}{2}$, letting $\lambda_0 = C_{\varepsilon_0}$, we get (6.16) for $u \in \mathscr{S}(\mathbb{R}^d)$. Now if $u \in \mathscr{C}^{\alpha+\beta}$, let $u_n := n^d\eta(n\cdot) * \big(\chi(\frac{\cdot}{n})u\big) \in \mathscr{S}(\mathbb{R}^d)$, where χ is the same function in section 2 and $\eta \in C_c^\infty(B_1)$, $\int \eta = 1$. $f_n := \lambda u_n - \mathscr{L}_\kappa^\alpha u_n - b \cdot \nabla u_n$. So

$$(\lambda - \lambda_0)\|u_n\|_{\mathscr{C}^\beta} + \|u_n\|_{\mathscr{C}^{\alpha+\beta}} \leqslant C\|f_n\|_{\mathscr{C}^\beta},$$

by this, we obtain

$$\begin{aligned}(\lambda - \lambda_0)\|u\|_{\mathscr{C}^\beta} + \|u\|_{\mathscr{C}^{\alpha+\beta}} \leqslant& \limsup_{n\to\infty}\Big[(\lambda - \lambda_0)\|u_n\|_{\mathscr{C}^\beta} + \|u_n\|_{\mathscr{C}^{\alpha+\beta}}\Big]\\ \leqslant& C\limsup_{n\to\infty}\|f_n\|_{\mathscr{C}^\beta} \leqslant C\|f\|_{\mathscr{C}^\beta}.\end{aligned}$$

For $\alpha \in (1,2)$, we only prove the case $\beta \leqslant 0$ here. By choosing $\gamma \in (-\beta, \frac{\alpha-1}{2})$, and Bony's decomposition, we have

$$\begin{aligned}&\|\Delta_j(b\cdot\nabla u)\|_\infty\\ =&\|\Delta_j(T_b\nabla u) + \Delta_j(T_{\nabla u}b) + \Delta_j(R(b,\nabla u))\|_\infty\\ \leqslant&\sum_{\substack{k\leqslant l-2;\\ |j-l|\leqslant 3}}\|\Delta_k b\|_\infty\|\Delta_l\nabla u\|_\infty + \sum_{\substack{l\leqslant k-2;\\ |j-k|\leqslant 3}}\|\Delta_k b\|_\infty\|\Delta_l\nabla u\|_\infty + \sum_{\substack{|k-l|\leqslant 1;\\ k,l\geqslant j-2}}\|\Delta_k b\|_\infty\|\Delta_l\nabla u\|_\infty\\ \leqslant& C_\gamma\|\nabla u\|_{\mathscr{C}^\gamma}\|b\|_{\mathscr{C}^\beta}\left(j2^{-\beta j}\cdot 2^{-\gamma j} + 2^{-\beta j} + 2^{-(\beta+\gamma)j}\right)\\ \leqslant& C_\gamma\|u\|_{\mathscr{C}^{1+\gamma}}\|b\|_{\mathscr{C}^\beta}2^{-\beta j}.\end{aligned}$$

Notice that,

$$\lambda\Delta_j u - \mathscr{L}_\kappa^\alpha\Delta_j u = -\Delta_j(b\cdot\nabla u) + \Delta_j f.$$

Like before, we have

$$\begin{aligned}(\lambda + c2^{\alpha j}))\|\Delta_j u\|_\infty =& \mathrm{sgn}(\Delta_j u(x_j))\cdot[\lambda\Delta_j u(x_j) + c2^{\alpha j}\Delta_j u(x_j)]\\ \leqslant&\|\lambda\Delta_j u - \mathscr{L}_\kappa^\alpha\Delta_j u\|_\infty\\ \leqslant&\|\Delta_j f\|_\infty + \|\Delta_j(b\cdot\nabla u)\|_\infty\\ \leqslant& C2^{-\beta j}(\|f\|_{\mathscr{C}^\beta} + \|b\|_{\mathscr{C}^\beta}\|u\|_{\mathscr{C}^{1+\gamma}}).\end{aligned}$$

Noticing that $1 + \gamma < \alpha + \beta$, by interpolation, we get (6.16). □

The next lemma will be used later.

Lemma 6.16. *Suppose $\kappa(z)$ satisfies ($\mathbf{H}_5$), then there is a constant $C = C(\alpha, d) > 0$ such that for all $\beta \in \mathbb{R}$ and $u \in \mathscr{C}^{\alpha+\beta}$,*

$$\|\mathscr{L}_\kappa^\alpha u\|_{\mathscr{C}^\beta} \leqslant C\Lambda_2\|f\|_{\mathscr{C}^{\alpha+\beta}}.$$

Proof. Recall that $\tilde{\varphi}$ is a smooth function supported in $B_2 \setminus B_{1/4}$ with $\tilde{\varphi} = 1$ on $B_{3/2} \setminus B_{1/2}$ and $\tilde{h} := \mathcal{F}^{-1}(\tilde{\varphi})$. Since $\tilde{h} \in \mathscr{S}$, it is easy to see that for some $c = c(\alpha, d) > 0$,

$$\|\mathscr{L}_\kappa^\alpha \tilde{h}\|_1 \leqslant C\Lambda_2 < \infty.$$

Let $\tilde{h}_j := \mathcal{F}^{-1}(\tilde{\varphi}(2^{-j}\cdot))$ for $j = 0, 1, 2, \cdots$. By scaling, we have

$$\|\mathscr{L}_\kappa^\alpha \tilde{h}_j\|_1 \leqslant C\Lambda_2 2^{\alpha j}, \quad j = 0, 1, 2, \cdots.$$

Since $\widehat{\Delta_j f} = \varphi(2^{-j}\cdot)\hat{f} = \tilde{\varphi}(2^{-j}\cdot)\varphi(2^{-j}\cdot)\hat{f}$, we have $\Delta_j f = \tilde{h}_j * \Delta_j f$ and

$$\|\Delta_j \mathscr{L}_\kappa^\alpha f\|_\infty = \|\mathscr{L}_\kappa^\alpha (\tilde{h}_j * (\Delta_j f))\|_\infty \leqslant \|\mathscr{L}_\kappa^\alpha \tilde{h}_j\|_1 \|\Delta_j f\|_\infty \leqslant C\Lambda_2 2^{\alpha j} \|\Delta_j f\|_\infty.$$

Similarly, one can show

$$\|\Delta_{-1} \mathscr{L}_\kappa^\alpha f\|_\infty \leqslant C\Lambda_2 \|\Delta_{-1} f\|_\infty.$$

Hence,

$$\|\mathscr{L}_\kappa^\alpha f\|_{\mathscr{C}^\beta} = \sup_{j \geqslant -1} 2^{\beta j} \|\Delta_j \mathscr{L}_\kappa^\alpha f\|_\infty \leqslant C\Lambda_2 \sup_{j \geqslant -1} 2^{\beta j} 2^{\alpha j} \|\Delta_j f\|_\infty = C\Lambda_2 \|f\|_{\mathscr{C}^{\alpha+\beta}}.$$

The proof is complete. □

6.3.2 The general case

Denote

$$\delta_z f(x) := f(x+z) - f(x), \quad \delta_z^\alpha f(x) := f(x+z) - f(x) - z^{(\alpha)} \cdot \nabla f(x).$$

We need the following lemma.

Lemma 6.17. *Suppose $\alpha \in (0, 2)$ and $\kappa(x, z)$ satisfies ($\mathbf{H}_2$) and ($\mathbf{H}_3$), then*

1. *for any $\beta \in (0, \vartheta]$, we have*

$$\|\mathscr{L}_\kappa^\alpha u\|_{\mathscr{C}^\beta} \leqslant C\Lambda_2 \|u\|_{\mathscr{C}^{\alpha+\beta}} + C_\theta \Lambda_3 \|u\|_{\mathscr{C}^{\alpha+\theta}}, \tag{6.17}$$

where $\theta \in (0, \beta)$.

2. *for any $\beta \in (-(\alpha \wedge \vartheta), 0]$, we have*

$$\|\mathscr{L}_\kappa^\alpha u\|_{\mathscr{C}^\beta} \leqslant C(\Lambda_2 + \Lambda_3) \|u\|_{\mathscr{C}^{\alpha+\beta}}. \tag{6.18}$$

Proof. (1). Suppose $\alpha \in (0, 1]$ and $\beta \in (0, \vartheta]$. For any $x_0 \in \mathbb{R}^d$, define

$$\mathscr{L}_0^\alpha u(x) = \int_{\mathbb{R}^d} \delta_z^\alpha u(x) \frac{\kappa(x_0, z)}{|z|^{d+\alpha}} \mathrm{d}z.$$

Notice that $|\mathscr{L}_\kappa^\alpha u(x_0)| = |\mathscr{L}_0^\alpha u(x_0)|$, by Lemma 6.16, we get

$$\|\mathscr{L}_0^\alpha u\|_{L^\infty} \leqslant C\Lambda_2 \|u\|_{\mathscr{C}^{\alpha+\beta}}.$$

For any $x \in B_1(x_0)$ and $\theta \in (0, \beta)$, by definition

$$
\begin{aligned}
|\mathscr{L}_\kappa^\alpha u(x) - \mathscr{L}_0^\alpha u(x)| \leqslant & \left| \int_{\mathbb{R}^d} \delta_z^\alpha u(x) \frac{(\kappa(x,z) - \kappa(x_0, z))}{|z|^{d+\alpha}} \mathrm{d}z \right| \\
\leqslant & \Lambda_3 |x - x_0|^\beta \int_{\mathbb{R}^d} |\delta_z^\alpha u(x)| \frac{\mathrm{d}z}{|z|^{d+\alpha}} \\
\leqslant & C_\theta \Lambda_3 |x - x_0|^\beta \|u\|_{C^{\alpha+\theta}}.
\end{aligned}
$$

Since

$$
|\mathscr{L}_\kappa^\alpha u(x) - \mathscr{L}_\kappa^\alpha u(x_0)| \leqslant |\mathscr{L}_\kappa^\alpha u(x) - \mathscr{L}_0^\alpha u(x)| + |\mathscr{L}_0^\alpha u(x) - \mathscr{L}_0^\alpha u(x_0)|,
$$

by the Lemma 6.16, if $\beta \in (0, \vartheta]$,

$$
|\mathscr{L}_0^\alpha u(x) - \mathscr{L}_\kappa^\alpha u(x_0)| \leqslant |\mathscr{L}_0^\alpha u(x) - \mathscr{L}_0^\alpha u(x_0)| \leqslant C\Lambda_2 \|u\|_{\mathscr{C}^{\alpha+\beta}} |x - x_0|^\beta.
$$

Combining the above inequalities, we get (6.17).

(2). We only prove the case $\alpha \in (1,2)$ and $\beta \in (-\vartheta, 0]$, which is harder and the only case that will be used below. Denote $\kappa_z(y) := \kappa(y,z)$, by definite we have

$$
\begin{aligned}
\Delta_j \mathscr{L}_\kappa^\alpha u(x) = & \int_{\mathbb{R}^d} h_j(x-y) \ \mathrm{d}y \int_{\mathbb{R}^d} \delta_z^\alpha u(y) \ \frac{\kappa(y,z)}{|z|^{d+\alpha}} \ \mathrm{d}z \\
= & \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \left[\delta_z^\alpha u(y) \kappa_z(y) \right] h_j(x-y) \ \mathrm{d}y \right) \ \frac{\mathrm{d}z}{|z|^{d+\alpha}}.
\end{aligned}
\tag{6.19}
$$

Denote

$$
I_j(x,z) = \int_{\mathbb{R}^d} \left[\delta_z^\alpha u(y) \kappa_z(y) \right] h_j(x-y) \ \mathrm{d}y,
$$

We drop the index x below for simple. By Bony's decomposition,

$$
\begin{aligned}
|I_j(z)| = & \left| \Delta_j \sum_{k,l \geqslant -1} \left[(\delta_z^\alpha \Delta_k u) \cdot \Delta_l \kappa_z \right] \right| \\
= & \left| \Delta_j \left(\sum_{k \leqslant l-2} \delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z + \sum_{l \leqslant k-2} \delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z + \sum_{|k-l| \leqslant 1} \delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z \right) \right| \\
\leqslant & \sum_{\substack{k \leqslant l-2; \\ |l-j| \leqslant 3}} \left| \delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z \right| + \sum_{\substack{l \leqslant k-2; \\ |k-j| \leqslant 3}} \left| \delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z \right| + \sum_{\substack{k,l \geqslant j-2; \\ |k-l| \leqslant 1}} \left| \delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z \right| \\
=: & I_j^{(1)}(z) + I_j^{(2)}(z) + I_j^{(3)}(z).
\end{aligned}
$$

Roughly speaking, the first inequality above holds because the Fourier transforms of $\sum_{k:k \leqslant l-2} \Delta_k f \Delta_l g$ and $\mathbf{1}_{|k-l| \leqslant 1} \Delta_k f \Delta_l g$ are supported around $2^l \mathcal{C}$ and $2^l B_1$ respectively. Noticing that by Bernstein's inequality

$$
\begin{aligned}
|\delta_z^\alpha \Delta_k u(y)| = & \left| \int_0^1 z \cdot \left[\nabla \Delta_k u(y+tz) - \nabla \Delta_k u(y) \right] \mathrm{d}t \right| \\
\leqslant & 2|z| \|\nabla \Delta_k u\|_\infty \leqslant C \|u\|_{\mathscr{C}^\gamma} |z| 2^{(1-\gamma)k},
\end{aligned}
\tag{6.20}
$$

and

$$|\delta_z^\alpha \Delta_k u(y)| \leqslant |z|^2 \|\nabla^2 \Delta_k u\|_\infty \leqslant C\|u\|_{\mathscr{C}^\gamma} |z|^2 2^{(2-\gamma)k}, \tag{6.21}$$

where $\gamma := \alpha + \beta$. Next we estimate each $I_j^{(i)}(z)$, we only need to care about the case when j is large, say $j \geqslant 10$.

- If $|z| < 2^{-j}$: for $I_j^{(1)}(z)$, by (6.21) and noticing that $2 - \gamma > 0$, we have

$$\begin{aligned} I_j^{(1)}(z) = &\sum_{\substack{k\leqslant l-2;\\ |l-j|\leqslant 3}} \left|\delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z\right| \\ \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} \sum_{k \lesssim j} |z|^2 2^{(2-\gamma)k} 2^{-\vartheta j} \\ \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} |z|^2 2^{(2-\gamma-\vartheta)j}. \end{aligned} \tag{6.22}$$

  Similarly, for $I_j^{(2)}(z)$, by (6.21) and noticing that $\vartheta > 0$, we have

$$\begin{aligned} I_j^{(2)}(z) \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} \sum_{l \lesssim j} |z|^2 2^{(2-\gamma)j} 2^{-\vartheta l} \\ \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} |z|^2 2^{(2-\gamma)j}. \end{aligned} \tag{6.23}$$

  For $I_j^{(3)}(z)$, we choose $\varepsilon_0 \in (0, (\beta+\vartheta) \wedge (2-\alpha))$, by (6.20), (6.21) and noticing that $1 - \gamma - \vartheta < 0$ and $2 - \alpha - \varepsilon_0 > 0 \vee (2 - \gamma - \vartheta)$, we have

$$\begin{aligned} I_j^{(3)}(z) \leqslant &\sum_{\substack{|k-l|\leqslant 1;\\ k,l > -\log_2|z| - 2}} \left|\delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z\right| + \sum_{\substack{|k-l|\leqslant 1;\\ j-3\leqslant k,l\leqslant -\log_2|z|}} \left|\delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z\right| \\ \leqslant &C\|\kappa_z\|_{\mathscr{C}^\vartheta} \|u\|_{\mathscr{C}^\gamma} \Big(\sum_{k \geqslant -\log_2|z|} |z| 2^{(1-\gamma)k} 2^{-\vartheta k} + \sum_{j-2\leqslant k \leqslant -\log_2|z|} |z|^2 2^{(2-\gamma)k} 2^{-\vartheta k} \Big) \\ \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} \left(|z|^{\gamma+\vartheta} + |z|^2 \sum_{k \leqslant -\log_2|z|} 2^{(2-\alpha-\varepsilon_0)k} \right) \\ \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} |z|^{\alpha+\varepsilon_0}. \end{aligned} \tag{6.24}$$

- $|z| \geqslant 2^{-j}$: for $I_j^{(1)}(z)$, notice that $\gamma < 2$ and $1 - \gamma - \vartheta < 0$, we have

$$\begin{aligned} I_j^{(1)}(z) \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} \Big(\sum_{-(1\wedge \log_2|z|)\leqslant k \leqslant j} |z| 2^{(1-\gamma)k} 2^{-\vartheta j} \\ &+ \sum_{-1\leqslant k < -(1\wedge\log_2|z|)} |z|^2 2^{(2-\gamma)k} 2^{-\vartheta j} \Big) \\ \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} \Big(\mathbf{1}_{\gamma>1} |z|^\gamma 2^{-\vartheta j} \\ &+ \mathbf{1}_{\gamma=1} |z| j 2^{-\vartheta j} + \mathbf{1}_{\gamma<1} |z| 2^{(1-\gamma-\vartheta)j} + |z|^\gamma 2^{-\vartheta j} \Big) \\ \leqslant &C\|u\|_{\mathscr{C}^\gamma} \|\kappa_z\|_{\mathscr{C}^\vartheta} (|z|^\gamma + \mathbf{1}_{\gamma<1} |z| 2^{(1-\gamma-\vartheta)j}). \end{aligned} \tag{6.25}$$

For $I_j^{(2)}(z)$, by (6.20) noticing that $\vartheta > 0$, we have

$$\begin{aligned} I_j^{(2)}(z) \leqslant & \sum_{\substack{l\leqslant k-2;\\ |k-j|\leqslant 3}} \left|(\delta_z^\alpha \Delta_k u \cdot \Delta_l \kappa_z)\right| \\ \leqslant & C\|u\|_{\mathscr{C}^\gamma}\|\kappa_z\|_{\mathscr{C}^\vartheta} \sum_{l\lesssim j} |z|2^{(1-\gamma)j}2^{-\vartheta l} \\ \leqslant & C\|u\|_{\mathscr{C}^\gamma}\|\kappa_z\|_{\mathscr{C}^\vartheta}|z|2^{(1-\gamma)j}. \end{aligned} \tag{6.26}$$

For $I_j^{(3)}(z)$, by (6.21), and notice that $1-\gamma-\vartheta<0$, we have

$$\begin{aligned} I_j^{(3)}(z) \leqslant & C\|u\|_{\mathscr{C}^\gamma}\|\kappa_z\|_{\mathscr{C}^\vartheta} \sum_{\substack{k,l\geqslant j-2;\\ |k-l|\leqslant 2}} |z|2^{(1-\gamma)k}2^{-\vartheta l} \\ \leqslant & C\|u\|_{\mathscr{C}^\gamma}\|\kappa_z\|_{\mathscr{C}^\vartheta}|z|2^{(1-\gamma-\vartheta)j}. \end{aligned} \tag{6.27}$$

Combining (6.22)-(6.27) and recalling that $\gamma=\alpha+\beta$, we obtain that for each $x\in\mathbb{R}^d$,

$$\begin{aligned} |\Delta_j \mathscr{L}_\kappa^\alpha u(x)| = & \left|\int_{\mathbb{R}^d} I_j(x,z)\frac{\mathrm{d}z}{|z|^{d+\alpha}}\right| \\ \leqslant & C\|u\|_{\mathscr{C}^\gamma}\sup_{z\in\mathbb{R}^d}\|\kappa_z\|_{\mathscr{C}^\vartheta}\Big(2^{(2-\gamma)j}\int_{|z|<2^{-j}}|z|^{2-d-\alpha}\mathrm{d}z + \int_{|z|<2^{-j}}|z|^{\varepsilon_0-d}\mathrm{d}z \\ & + \int_{|z|\geqslant 2^{-j}}|z|^{\gamma-d-\alpha}\mathrm{d}z + 2^{(1-\gamma)j}\int_{|z|\geqslant 2^{-j}}|z|^{1-d-\alpha}\mathrm{d}z\Big) \\ = & C\|u\|_{\mathscr{C}^\gamma}\sup_{z\in\mathbb{R}^d}\|\kappa_z\|_{\mathscr{C}^\vartheta}\Big(2^{-\beta j}+1+2^{-\beta j}\Big) \\ \leqslant & C\|u\|_{\mathscr{C}^\gamma}\sup_{z\in\mathbb{R}^d}\|\kappa_z\|_{\mathscr{C}^\vartheta}2^{-\beta j}. \end{aligned} \tag{6.28}$$

i.e.

$$\|\mathscr{L}_\kappa^\alpha u\|_{\mathscr{C}^\beta} = \sup_{j\geqslant -1} 2^{-\beta j}\|\Delta_j\mathscr{L}_\kappa^\alpha u\|_\infty \leqslant C\|u\|_{\mathscr{C}^{\alpha+\beta}}\sup_{z\in\mathbb{R}^d}\|\kappa_z\|_{\mathscr{C}^\vartheta}.$$

So we complete our proof. □

Before we proving our main results, let us give a brief discussion about our assumptions on $\kappa(x,z)$: let$\Lambda=\Lambda_1/(2c_d)$, where c_d is the volume of unity ball in $\mathbb{R}^d$. By our assumptions (**H**$_1$) and (**H**$_2$), we can see that for any $r\in(0,r_0]$, $x\in\mathbb{R}^d$,

$$\begin{aligned} |B_r\cap\{\kappa(x,\cdot)\geqslant\Lambda\}| \geqslant & \Lambda_2^{-1}\int_{B_r\cap\{\kappa(x,\cdot)\geqslant\Lambda\}}\kappa(x,z)\mathrm{d}z \\ = & \Lambda_2^{-1}\int_{B_r}\kappa(x,z)\mathrm{d}z - \Lambda_2^{-1}\int_{B_r\cap\{\kappa(x,\cdot)<\Lambda\}}\kappa(x,z)\mathrm{d}z \\ \geqslant & \Lambda_2^{-1}(\Lambda_1 r^d - \Lambda|B_r|) \geqslant \frac{\Lambda_1}{2\Lambda_2}r^d. \end{aligned}$$

Thus, for each $x\in\mathbb{R}^d$, $\kappa(x,\cdot)$ satisfies (**H**$_4$) with $\Lambda=\Lambda_1/(2c_d)$ and $\delta_0=\Lambda_1/(2\Lambda_2)$.

Now we give the proof for Theorem 6.1.

Proof of Theorem 6.1. (1) Define

$$\mathscr{L}_0^\alpha u(x) = \int_{\mathbb{R}^d} \delta_z^\alpha f(x) \frac{\kappa(x_0, z)}{|z|^{d+\alpha}} \mathrm{d}z.$$

Choose η be a smooth function with compact support in B_1 and $\eta(x) = 1$, if $x \in B_{\frac{1}{2}}$. Fixed $x_0 \in \mathbb{R}^d$, define

$$\eta_\varepsilon^{x_0}(x) := \eta\Big(\frac{x - x_0}{\varepsilon}\Big); \quad \kappa_\varepsilon^{x_0}(x, z) := [\kappa(x, z) - \kappa(x_0, z)]\eta_\varepsilon(x).$$

We omit the supscript x_0 below for simple. Define $v = u\eta_\varepsilon$, then we have

$$\begin{aligned} &\lambda v - \mathscr{L}_0^\alpha v - b \cdot \nabla v \\ =&[\eta_\varepsilon f - ub \cdot \nabla\eta_\varepsilon + u\mathscr{L}_\kappa^\alpha \eta_\varepsilon] + \eta_\varepsilon(\mathscr{L}_\kappa^\alpha u - \mathscr{L}_0^\alpha u) + [\eta_\varepsilon \mathscr{L}_0^\alpha u - \mathscr{L}_0^\alpha(\eta_\varepsilon u) - u\mathscr{L}_\kappa^\alpha \eta_\varepsilon]. \end{aligned} \tag{6.29}$$

Obviously,

$$\|\eta_\varepsilon f - ub \cdot \nabla\eta_\varepsilon + u\mathscr{L}_\kappa^\alpha \eta_\varepsilon\|_{\mathscr{C}^\beta} \leqslant C_\varepsilon(\|f\|_{\mathscr{C}^\beta} + \|u\|_{\mathscr{C}^\beta}). \tag{6.30}$$

Denote

$$\widetilde{w}_\varepsilon(x) := \eta_\varepsilon(x)(\mathscr{L}_\kappa^\alpha u(x) - \mathscr{L}_0^\alpha u(x)) = \int_{\mathbb{R}^d} \delta_z^\alpha u \frac{\kappa_\varepsilon(x, z)}{|z|^{d+\alpha}} \mathrm{d}z.$$

By (6.17), for any $\theta \in (0, \beta)$

$$\begin{aligned} \|\widetilde{w}_\varepsilon\|_{\mathscr{C}^\beta} \leqslant& C \sup_z \|\kappa_\varepsilon(\cdot, z)\|_{L^\infty} \|u\|_{\mathscr{C}^{\alpha+\beta}} + C_\theta \sup_z [\kappa_\varepsilon(\cdot, z)]_\vartheta \|u\|_{\mathscr{C}^{\alpha+\theta}} \\ \leqslant& C\varepsilon^\beta \|u\|_{\mathscr{C}^{\alpha+\beta}} + C_{\theta,\varepsilon}\|u\|_{\mathscr{C}^{\alpha+\theta}}. \end{aligned} \tag{6.31}$$

Denote

$$w_\varepsilon(x) := [\eta_\varepsilon \mathscr{L}_0^\alpha u - \mathscr{L}_0^\alpha(\eta_\varepsilon u) - u\mathscr{L}_0^\alpha \eta_\varepsilon](x)$$

and $\delta_z f(x) = (f(x + z) - f(x))$, by definition, we have

$$w_\varepsilon(x) = \int_{\mathbb{R}^d} \delta_z \eta_\varepsilon(x) \ \delta_z u(x) \frac{\kappa(x_0, z)}{|z|^{d+\alpha}} \mathrm{d}z, \tag{6.32}$$

and

$$\begin{aligned} w_\varepsilon(x) - w_\varepsilon(y) =& \int_{\mathbb{R}^d} \delta_z \eta_\varepsilon(x) \big[\delta_z u(x) - \delta_z u(y)\big] \frac{\kappa(x_0, z)}{|z|^{d+\alpha}} \mathrm{d}z \\ &+ \int_{\mathbb{R}^d} \big[\delta_z \eta_\varepsilon(x) - \delta_z \eta_\varepsilon(y)\big] \delta_z u(y) \frac{\kappa(x_0, z)}{|z|^{d+\alpha}} \mathrm{d}z. \end{aligned} \tag{6.33}$$

In order to estimate the $\mathscr{C}^\beta$ norm of w_ε, for different cases we have to deal it separately.

(i)For $\alpha \in (0, 1)$, by (6.32),

$$\begin{aligned} |w_\varepsilon(x)| \leqslant& \int_{|z|\leqslant 1} \|\nabla\eta_\varepsilon\|_{L^\infty} \|u\|_{L^\infty} |z| \frac{\kappa(x_0, z)}{|z|^{d+\alpha}} \mathrm{d}z + 2\int_{|z|>1} \|\eta_\varepsilon\|_{L^\infty} \|u\|_{L^\infty} \frac{\kappa(x_0, z)}{|z|^{d+\alpha}} \mathrm{d}z \\ \leqslant& C_\varepsilon \|u\|_{L^\infty}. \end{aligned}$$

And by (6.33),

$$
\begin{aligned}
&\big|w_\varepsilon(x)-w_\varepsilon(y)\big|\\
\leqslant& C|x-y|^\beta\Big(\|u\|_{C^\beta}\|\nabla\eta_\varepsilon\|_{L^\infty}\int_{|z|\leqslant 1}\frac{\mathrm{d}z}{|z|^{d+\alpha-1}}+\|u\|_{C^\beta}\|\eta_\varepsilon\|_{L^\infty}\int_{|z|>1}\frac{\mathrm{d}z}{|z|^{d+\alpha}}\Big)\\
&+C|x-y|^\beta\Big(\|\eta_\varepsilon\|_{\mathscr{C}^{1+\beta}}\|u\|_{L^\infty}\int_{|z|\leqslant 1}\frac{\mathrm{d}z}{|z|^{d+\alpha-1}}+\|\eta_\varepsilon\|_{C^\beta}\|u\|_{L^\infty}\int_{|z|>1}\frac{\mathrm{d}z}{|z|^{d+\alpha}}\Big)\\
\leqslant& C_\varepsilon|x-y|^\beta\|u\|_{\mathscr{C}^\beta}.
\end{aligned}
$$

Hence, we have

$$
\|w_\varepsilon\|_{\mathscr{C}^\beta}\leqslant C_\varepsilon\|u\|_{\mathscr{C}^\beta}. \tag{6.34}
$$

Let λ_0' be the constant λ_0 in Theorem 6.15, by (6.29), (6.30), (6.31), (6.34), Theorem 6.15, interpolation theorem and the discussion before this proof, we have

$$
\begin{aligned}
&\|u\|_{C^{\alpha+\beta}(B_{\varepsilon/2}(x_0))}+(\lambda-\lambda_0')\|u\eta_\varepsilon^{x_0}\|_{\mathscr{C}^\beta}\\
\leqslant& C\|v\|_{\mathscr{C}^{\alpha+\beta}}+(\lambda-\lambda_0')\|v\|_{\mathscr{C}^\beta}\\
\leqslant& C\varepsilon^\beta\|u\|_{\mathscr{C}^{\alpha+\beta}}+C_{\theta,\varepsilon}\|u\|_{\mathscr{C}^{\alpha+\theta}}+C\|f\|_{\mathscr{C}^\beta}\\
\leqslant& C\varepsilon^\beta\|u\|_{C^{\alpha+\beta}}+C_{\theta,\varepsilon}\|u\|_{\mathscr{C}^\beta}+C\|f\|_{\mathscr{C}^\beta}\\
\leqslant& C\varepsilon^\beta\sup_{x_0\in\mathbb{R}^d}\|u\|_{C^{\alpha+\beta}(B_{\varepsilon/2}(x_0))}+C_{\theta,\varepsilon}\|u\|_{\mathscr{C}^\beta}+C\|f\|_{\mathscr{C}^\beta}.
\end{aligned}
$$

We can fixed ε_0 sufficiently small, such that $C\varepsilon_0^\beta\leqslant 1/2$, so we have

$$
\sup_{x_0\in\mathbb{R}^d}\Big(\|u\|_{C^{\alpha+\beta}(B_{\varepsilon_0/2}(x_0))}+(\lambda-\lambda_0')\|u\eta_\varepsilon^{x_0}\|_{\mathscr{C}^\beta}\Big)\leqslant C_{\varepsilon_0}(\|f\|_{\mathscr{C}^\beta}+\|u\|_{\mathscr{C}^\beta}).
$$

This yields

$$
\|u\|_{\mathscr{C}^{\alpha+\beta}}\leqslant C_{\varepsilon_0}\sup_{x_0\in\mathbb{R}^d}\|u\|_{C^{\alpha+\beta}(B_{\varepsilon_0/2}(x_0))}\leqslant C_{\varepsilon_0}\left(\|u\|_{\mathscr{C}^\beta}+\|f\|_{\mathscr{C}^\beta}\right),
$$

and

$$
C_{\varepsilon_0}(\|f\|_{\mathscr{C}^\beta}+\|u\|_{\mathscr{C}^\beta})\geqslant(\lambda-\lambda_0')\sup_{x_0\in\mathbb{R}^d}\|u\eta_{\varepsilon_0}^{x_0}\|_{\mathscr{C}^\beta}\geqslant c_{\varepsilon_0}(\lambda-\lambda_0')\|u\|_{\mathscr{C}^\beta},
$$

where c_{ε_0} is a constant larger than 0. Thus,

$$
\|u\|_{\mathscr{C}^{\alpha+\beta}}+(\lambda-\lambda_0')\|u\|_{\mathscr{C}^\beta}\leqslant C_{\epsilon_0}(\|f\|_{\mathscr{C}^\beta}+\|u\|_{\mathscr{C}^\beta}).
$$

Letting $\lambda_0=\lambda_0'+C_{\varepsilon_0}$, we obtain (6.2).

(ii)For $\alpha=1$, by (6.32) and (6.33), we have

$$
\|w_\varepsilon\|_{L^\infty}\leqslant C_\varepsilon\|u\|_{C^1},
$$

and

$$
\begin{aligned}
&|w_\varepsilon(x)-w_\varepsilon(y)|\\
\leqslant&\left|\int_{|z|\leqslant\delta}\delta_z\eta_\varepsilon(x)\Big[\delta_z u(x)-\delta_z u(y)\Big]\frac{\kappa(x_0,z)}{|z|^{d+\alpha}}\,\mathrm{d}z\right|+\left|\int_{|z|>\delta}\delta_z\eta_\varepsilon(x)\Big[\delta_z u(x)-\delta_z u(y)\Big]\frac{\kappa(x_0,z)}{|z|^{d+\alpha}}\,\mathrm{d}z\right|\\
+&\left|\int_{|z|\leqslant\delta}\delta_z u(y)\Big[\delta_z\eta_\varepsilon(x)-\delta_z\eta_\varepsilon(y)\Big]\frac{\kappa(x_0,z)}{|z|^{d+\alpha}}\,\mathrm{d}z\right|+\left|\int_{|z|>\delta}\delta_z u(y)\Big[\delta_z\eta_\varepsilon(x)-\delta_z\eta_\varepsilon(y)\Big]\frac{\kappa(x_0,z)}{|z|^{d+\alpha}}\,\mathrm{d}z\right|\\
\leqslant& C\varepsilon^{-1}|x-y|^\beta\|u\|_{\mathscr{C}^{1+\beta}}\int_{|z|\leqslant\delta}\frac{\mathrm{d}z}{|z|^{d+\alpha-2}}+C|x-y|^\beta\|u\|_{C^\beta}\int_{|z|>\delta}\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\
&+C\varepsilon^{-1-\beta}|x-y|^\beta\|\nabla u\|_{L^\infty}\int_{|z|\leqslant\delta}\frac{\mathrm{d}z}{|z|^{d+\alpha-2}}\mathrm{d}z+C\varepsilon^{-\beta}|x-y|^\beta\|u\|_{L^\infty}\int_{|z|>\delta}\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\
\leqslant&\Big(C\varepsilon^{-2}\delta^{2-\alpha}\|u\|_{\mathscr{C}^{1+\beta}}+C(\varepsilon,\delta)\|u\|_{\mathscr{C}^\beta}\Big)|x-y|^\beta.
\end{aligned}
\tag{6.35}
$$

Hence,

$$\|w_\varepsilon\|_{\mathscr{C}^\beta}\leqslant C\varepsilon^{-2}\delta^{2-\alpha}\|u\|_{\mathscr{C}^{1+\beta}}+C(\varepsilon,\delta)\|u\|_{\mathscr{C}^\beta}.$$

Choosing $\delta=\varepsilon^{\frac{2+\beta}{2-\alpha}}$, by Theorem 6.15, interpolation and above inequality, we get

$$
\begin{aligned}
&\|u\|_{\mathscr{C}^{1+\beta}(B_{\varepsilon/2}(x_0))}+(\lambda-\lambda_0')\|u\|_{\mathscr{C}^\beta(B_{\varepsilon/2}(x_0))}\\
\leqslant& C\varepsilon^\beta\|u\|_{\mathscr{C}^{\alpha+\beta}}+C_{\theta,\varepsilon}\|u\|_{\mathscr{C}^{1+\theta}}+C\varepsilon^{-2}\delta^{2-\alpha}\|u\|_{\mathscr{C}^{1+\beta}}+C(\varepsilon,\delta)\|u\|_{\mathscr{C}^\beta}+C\|f\|_{\mathscr{C}^\beta}\\
\leqslant& C\varepsilon^\beta\|u\|_{\mathscr{C}^{1+\beta}}+C_\varepsilon\|u\|_{\mathscr{C}^\beta}+C\|f\|_{\mathscr{C}^\beta}.
\end{aligned}
$$

Like the above case, we get (6.2).

(2) For $\alpha\in(1,2)$, we only give the proof for $\beta\leqslant 0$ here. Like the previous cases, we have (6.29). Moreover, notice that $\beta\in(-(\frac{\alpha-1}{2}\wedge\vartheta),0]$, it is easy to see that

$$\|\eta_\varepsilon f-ub\cdot\nabla\eta_\varepsilon+u\mathscr{L}^\alpha_\kappa\eta_\varepsilon\|_{\mathscr{C}^\beta}\leqslant C_\varepsilon(\|f\|_{\mathscr{C}^\beta}+\|u\|_{\mathscr{C}^\vartheta}),$$

and

$$\|w_\varepsilon\|_{\mathscr{C}^\beta}\leqslant C\|w_\varepsilon\|_{L^\infty}\leqslant C\|u\|_{C^1}.$$

For $\tilde w_\varepsilon$, fixing $\gamma\in(-\beta,\vartheta)$, then for any $z\in\mathbb{R}^d$,

$$
\begin{aligned}
\|\kappa_\varepsilon(\cdot,z)\|_{\mathscr{C}^\gamma}=&\|[\kappa(\cdot,z)-\kappa(x_0,z)]\eta_\varepsilon(\cdot)\|_{\mathscr{C}^\gamma}\\
\leqslant& C(\varepsilon^\vartheta[\eta_\varepsilon]_\gamma+[\kappa(\cdot,z)]_{C^\gamma(B_\varepsilon(x_0))})\\
\leqslant& C\varepsilon^{\vartheta-\gamma}.
\end{aligned}
$$

Using Lemma 6.17 (2)(replace ϑ with γ) and above inequality, we obtain

$$
\begin{aligned}
\|\tilde w_\varepsilon\|_{\mathscr{C}^\beta}:=&\|\eta_\varepsilon(\mathscr{L}^\alpha_\kappa u-\mathscr{L}^\alpha_0 u)\|_{\mathscr{C}^\beta}=\left\|\int_{\mathbb{R}^d}\delta_z u(\cdot)\frac{\kappa_\varepsilon(\cdot,z)}{|z|^{d+\alpha}}\mathrm{d}z\right\|_{\mathscr{C}^\beta}\\
\leqslant& C\varepsilon^{\vartheta-\gamma}\|u\|_{\mathscr{C}^{\alpha+\beta}}.
\end{aligned}
$$

Now by the similar argument as in the previous case, we get (6.2).

□

6.4 Martingale solutions and weak solutions

Before going to the definition of martingale problem associated with $\mathscr{L}^{\alpha}_{\kappa,b}$, let us briefly introduce the corresponding SDE.

Let $(\Omega, \mathbf{P}, \mathcal{F})$ be a probability space and $N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s)$ be a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+$ with intensity measure is $\mathrm{d}r \, \frac{\mathrm{d}z}{|z|^{d+\alpha}} \, \mathrm{d}s$. Define

$$N^{(\alpha)}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) = \begin{cases} N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) & \alpha \in (0,1) \\ N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) - \mathrm{d}r \dfrac{\mathbf{1}_{B_1}(z)\mathrm{d}z}{|z|^{d+\alpha}}\mathrm{d}s & \alpha = 1 \\ N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) - \mathrm{d}r \dfrac{\mathrm{d}z}{|z|^{d+\alpha}}\mathrm{d}s & \alpha \in (1,2) \end{cases} \tag{6.36}$$

Consider the following SDE driven by Poisson random measure N:

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}^d} \int_0^{\infty} z \mathbf{1}_{[0, \kappa(X_{s-}, z))}(r) N^{(\alpha)}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) + \int_0^t b(X_s)\mathrm{d}s. \tag{6.37}$$

As mentioned before, when b is just a distribution, the drift term "$\int_0^{\cdot} b(X_s)\mathrm{d}s$" may not be a process with finite variation any more but an additive functional of X with zero energy, which means X may not be a semimartingale but a Dirichlet process. We give the precious definitions of Dirichlet processes and process of zero energy first.

Definition 6.18. *We say that a continuous adapted process $(A_t)_{t\in[0,T]}$ is a process of zero energy if $A_0 = 0$ and*

$$\lim_{\delta \to 0} \sup_{|\pi_T| < \delta} \mathbf{E}\Big(\sum_{t_i \in \pi_T} |A_{t_{i+1}} - A_{t_i}|^2 \Big) = 0$$

where π_T denotes a finite partition of $[0,T]$ and $|\pi_T|$ denotes the mesh size of the partition.

Definition 6.19. *We say that an adapted process $(X_t)_{t\in[0,T]}$ is a Dirichlet process if*

$$X_t = M_t + A_t \tag{6.38}$$

where M is a square-integrable martingale and A is an adapted process of zero energy.

Suppose $\kappa(\cdot, z), b$ is smooth and bounded, then the above equation has a unique solution. By Itô's formula(see [2, Theorem 4.4.7]), for any $f \in C_b^2$, we have

$$\begin{aligned} & f(X_t) - f(X_0) \\ = & \int_0^t \int_{\mathbb{R}^d} \int_0^{\infty} [f(X_{s-} + z\mathbf{1}_{[0,\kappa(X_{s-},z))}(r)) - f(X_{s-})] \widetilde{N}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) + \int_0^t b \cdot \nabla f(X_s)\mathrm{d}s \\ & + \int_0^t \int_{\mathbb{R}^d} \int_0^{\infty} [f(X_{s-} + z\mathbf{1}_{[0,\kappa(X_{s-},z))}(r)) - f(X_{s-}) - z\mathbf{1}_{[0,\kappa(X_{s-},z))}(r)] \mathrm{d}r \frac{\mathrm{d}z}{|z|^{d+\alpha}} \mathrm{d}s \\ = & M_t^f + \int_0^t \mathscr{L}^{\alpha}_{\kappa} f(X_s)\mathrm{d}s + \int_0^t b \cdot \nabla f(X_s)\mathrm{d}s, \end{aligned} \tag{6.39}$$

where

$$M_t^f := \int_0^t \int_{\mathbb{R}^d} \int_0^\infty [f(X_{s-} + z\mathbf{1}_{[0,\kappa(X_{s-},z))}(r)) - f(X_{s-})]\widetilde{N}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s).$$

Thus, (6.37) is the SDE associated with operator $\mathscr{L}_{\kappa,b}^{\alpha}$ at least when the coefficients are regular. However, when $b \in \mathscr{C}^\beta$ with $\beta \leqslant 0$, we must face up to the problem of how to define the term "$\int_0^t b(X_s)\mathrm{d}s$" in (6.37) and "$\int_0^t b \cdot \nabla f(X_s)\mathrm{d}s$" in (6.39). Inspired by [76], when considering the martingale problem associated with $\mathscr{L}_{\kappa,b}^{\alpha}$, if $b \in \mathscr{C}^\beta$ with $\beta \leqslant 0$, we need restrict ourselves to some probability measures on $\mathbb{D} := D(\mathbb{R}_+; \mathbb{R}^d)$ satisfying the following Krylov's type estimate:

Definition 6.20. *(Krylov's type estimate) We call a probability measure $\mathbb{P} \in \mathscr{P}(\mathbb{D})$ satisfy Krylov's estimate with indices μ if for any $T > 0$, there are positive constants C_T and γ such that for all $f \in C^\infty$, $0 \leqslant t_0 < t_1 \leqslant T$,*

$$\mathbb{E}\left|\int_{t_0}^{t_1} f(w_s)\mathrm{d}s\right|^2 \leqslant C_T|t_1 - t_0|^{1+\gamma}\|f\|_{\mathscr{C}^\mu}^2, \tag{6.40}$$

where the expectation $\mathbb{E}$ is taken with respect to $\mathbb{P}$. All the probability measure $\mathbb{P}$ with property (6.40) is denoted by $\mathscr{K}^\mu(\mathbb{D})$.

We should point out that for arbitrary $f \in \mathscr{C}^\beta$, there is no good smooth approximation sequence in space $\mathscr{C}^\beta$. However, the modifying approximation sequence $f_n := f * \eta_n$ converges to f in $\mathscr{C}^\mu$, for any $\mu < \beta$. So given $f \in \mathscr{C}^\beta$ with $\beta \leqslant 0$, in order to give a natural definition of $\int_0^t f(\omega_s)\mathrm{d}s$ under some suitable probability measure $\mathbb{P}$, we have to restrict ourselves to $\mathbb{P} \in \mathscr{K}^\mu(\mathbb{D})$ with $\mu < \beta$.

Proposition 6.21. *Let $\mu < \beta \leqslant 0$, $\mathbb{P} \in \mathscr{K}^\mu(\mathbb{D})$, for any $f \in \mathscr{C}^\beta$, there is a continuous $\mathcal{B}_t(\mathbb{D})$-adapted process A_t^f with zero energy and such that for any $T > 0$,*

$$\lim_{n\to\infty} \mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t f_n(w_s)\mathrm{d}s - A_t^f\right|\right) = 0, \tag{6.41}$$

where $C_b^\infty \ni f_n \xrightarrow{\mathscr{C}^\mu} f$. Moreover, the mapping $\mathscr{C}^\mu \ni f \mapsto A_\cdot^f \in L^2(\mathbb{D}, \mathbb{P}; C([0,T]))$ is a bounded linear operator and for all $0 \leqslant t_0 < t_1 \leqslant T$,

$$\mathbb{E}\left|A_{t_1}^f - A_{t_0}^f\right|^2 \leqslant C_T(t_1 - t_0)^{1+\gamma}\|f\|_{\mathscr{C}^\mu}^2, \tag{6.42}$$

where the constants C_T and γ are the same as in (6.40).

Since the proof for this proposition is just the same with Proposition 3.2 in [76], we omit the details here.

Now we are on the position to give the definition of martingale problem.

Definition 6.22 (Martingale Problem). *1. If $b \in B_b(\mathbb{R}^d)$, we call a probability measure $\mathbb{P} \in \mathscr{P}(\mathbb{D})$ a martingale solution associated with $\mathscr{L}^\alpha_{\kappa,b}$ starting from $x \in \mathbb{R}^d$ if for any $f \in C_b^\infty$,*

$$M_t^f := f(\omega_t) - f(x) - \int_0^t \mathscr{L}^\alpha_{\kappa,b} f(\omega_s) \mathrm{d}s \tag{6.43}$$

is a continuous $\mathcal{B}_t(\mathbb{D})$-martingale with $M_0^f = 0$ under $\mathbb{P}$. The set of the martingale solutions with starting point x is denoted by $\mathscr{M}_{\kappa,b}(x)$.

2. If $\mu < \beta \leqslant 0$, $b \in \mathscr{C}^\beta$ with $\beta \leqslant 0$, we call a probability measure $\mathbb{P} \in \mathscr{K}^\mu(\mathbb{D})$ a martingale solution associated with $\mathscr{L}^\alpha_{\kappa,b}$ starting from $x \in \mathbb{R}^d$ if for any $f \in C_b^\infty$,

$$M_t^f := f(w_t) - f(x) - \int_0^t \mathscr{L}^\alpha_{\kappa} f(w_s) \mathrm{d}s - A_t^{b\cdot\nabla f} \tag{6.44}$$

is a continuous $\mathcal{B}_t(\mathbb{D})$-martingale with $M_0^f = 0$ under $\mathbb{P}$. The set of the martingale solutions $\mathbb{P} \in \mathscr{K}^\mu(\mathbb{D})$ and starting point x is denoted by $\mathscr{M}^\mu_{\kappa,b}(x)$.

By Theorem 6.1 (1), immediately, we have

Lemma 6.23. *Suppose $\alpha \in (0,1]$, $\kappa(x,z)$ satisfies $(\mathbf{H}_1)$-$(\mathbf{H}_3)$ with $\max\{0,(1-\alpha)\} < \vartheta < 1$, and $b \in \mathscr{C}^\beta$ with $\beta \in (0,\vartheta)$, then for any $x \in \mathbb{R}^d$, there is a unique element in $\mathscr{M}_{\kappa,b}(x)$.*

Proof. The Existence of martingale solution to (6.43) is trivial, since the coefficients are globally Hölder continuous. We only give the proof for uniqueness. Suppose $\mathbb{P}_x \in \mathscr{M}_{\kappa,b}(x)$. For any $f \in C_b^\infty$ and $\lambda \geqslant \lambda_0$, where λ_0 is the constant in Theorem 6.1, let u be the solution to (6.1) and $u_n := u * \eta_n = n^d u * \eta(n\cdot)$. By the definition of $\mathbb{P}_x$ and Itô's formula, we have

$$\mathrm{e}^{-\lambda t} u_n(\omega_t) - u_n(\omega_0) = \int_0^t \mathrm{e}^{-\lambda s}[-\lambda u_n(\omega_s) + \mathscr{L}^\alpha_{\kappa,b} u_n(\omega_s)]\mathrm{d}s + \int_0^t \mathrm{e}^{-\lambda s} \mathrm{d}M_s^{u_n},$$

which implies

$$u_n(x) = \mathbb{E}_x\left(\int_0^\infty \mathrm{e}^{-\lambda t}[(\lambda u_n - \mathscr{L}^\alpha_{\kappa,b} u_n)(\omega_t)]\mathrm{d}s\right) = \mathbb{E}_x\left(\int_0^\infty \mathrm{e}^{-\lambda t} g_n(\omega_t)\mathrm{d}t\right), \tag{6.45}$$

where

$$g_n = f * \eta_n + [(\mathscr{L}^\alpha_\kappa u) * \eta_n - \mathscr{L}^\alpha_\kappa(u * \eta_n)] + [(b\cdot\nabla u) * \eta_n - b \cdot \nabla(u * \eta_n)]. \tag{6.46}$$

Noticing that $u \in \mathscr{C}^{\alpha+\beta}$ with $\beta > 0$, we have

$$\begin{aligned}
&[(\mathscr{L}^\alpha_\kappa u) * \eta_n - \mathscr{L}^\alpha_\kappa(u * \eta_n)](x) \\
=& \int_{\mathbb{R}^d} \eta_n(x-y)\mathrm{d}y \int_{\mathbb{R}^d} \delta_z^\alpha u(y) \frac{(\kappa(y,z) - \kappa(x,z))}{|z|^{d+\alpha}}\mathrm{d}z \\
\leqslant& \Lambda_3 \int_{\mathbb{R}^d} \eta_n(x-y)|x-y|^\vartheta \mathrm{d}y \int_{\mathbb{R}^d} \frac{|\delta_z^\alpha u(y)|}{|z|^{d+\alpha}}\mathrm{d}z \leqslant Cn^{-\vartheta}\|u\|_{\mathscr{C}^{\alpha+\beta}} \to 0 \ (n \to \infty).
\end{aligned}$$

And also $[(b\cdot\nabla u)*\eta_n - b\cdot\nabla(u*\eta_n)] \to 0$ uniformly in n. Hence, $\{g_n\}$ is uniformly bounded and converges to f. Taking limit in both side of (6.45), we obtain

$$u(x) = \mathbb{E}_x\left(\int_0^\infty \mathrm{e}^{-\lambda t} f(\omega_t)\mathrm{d}t\right),$$

which implies the one dimensional distribution of $\mathbb{P}_x$ is unique and thus the uniqueness of $\mathbb{P}_x$ follows(see [20] for details). □

Next we consider the case when $\alpha\in(1,2)$ and b is just a distribution.

Lemma 6.24. *Suppose $\alpha\in(1,2)$, $\kappa(x,z)$ satisfies $(\mathbf{H}_1)$-$(\mathbf{H}_3)$ and $b\in\mathscr{C}^\beta$ with $\beta\in(-(\frac{\alpha-1}{2}\wedge\vartheta),0]$. Then for each $x\in\mathbb{R}^d$, there is a unique probability measure $\mathbb{P}_x\in\mathscr{M}^\mu_{\kappa,b}(x)$, for some $\mu<\beta$.*

Proof. **Uniqueness:** The proof is similar with the one of Lemma 6.23. Suppose $-\vartheta<\mu<\beta$, $\mathbb{P}_x\in\mathscr{M}^\mu_{\kappa,b}(x)$, thanks to the fact $\mathbb{P}_x\in\mathscr{K}^\mu(\mathbb{D})$, we only need to show $g_n\to f$ in $\mathscr{C}^\mu$, where g_n is defined in (6.46). Notice that $\mathscr{L}^\alpha_\kappa u\in\mathscr{C}^\beta$, $u*\eta_n \overset{\mathscr{C}^{\alpha+\mu}}{\longrightarrow} u$ and by Lemma 6.17 $\mathscr{L}^\alpha_\kappa:\mathscr{C}^{\alpha+\mu}\to\mathscr{C}^\mu$ is bounded, we get

$$\begin{aligned}&\|(\mathscr{L}^\alpha_\kappa u)*\eta_n - \mathscr{L}^\alpha_\kappa(u*\eta_n)\|_{\mathscr{C}^\mu}\\ \leqslant&\|(\mathscr{L}^\alpha_\kappa u)*\eta_n - \mathscr{L}^\alpha_\kappa u\|_{\mathscr{C}^\mu} + \|\mathscr{L}^\alpha_\kappa(u*\eta_n) - \mathscr{L}^\alpha_\kappa u\|_{\mathscr{C}^\mu}\to 0,\quad (n\to\infty).\end{aligned}$$

Similarly, we have

$$\|(b\cdot\nabla u)*\eta_n - b\cdot\nabla(u*\eta_n)\|_{\mathscr{C}^\mu}\to 0,\quad (n\to\infty).$$

Thus we get $\lim_{n\to\infty}\|g_n-f\|_{\mathscr{C}^\mu} = \|f*\eta_n - f\|_{\mathscr{C}^\mu} = 0$.

Existence: Let $b_n = b*\eta_n$, $\kappa_n(\cdot,z) = (\kappa(\cdot,z)*\eta_n)(\cdot)$. Let X^n_t be the unique solution to the following SDE:

$$X^n_t = x + \int_0^t\int_{\mathbb{R}^d}\int_0^\infty z\mathbf{1}_{[0,\kappa_n(X^n_{s-},z))}\widetilde{N}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) + \int_0^t b_n(X^n_s)\mathrm{d}s,$$

where N and $\widetilde{N} = N^\alpha$ are defined at the beginning of this section. Then the probability measure $\mathbb{P}^n_x = \mathbf{P}\circ(X^n_t)^{-1}$ on $\mathbb{D}$ is an element in $\mathscr{M}_{\kappa_n,b_n}(x)$. For any $f\in C_b^\infty$, let u^λ_n be the solution to

$$\lambda u^\lambda_n - \mathscr{L}^\alpha_{\kappa_n,b_n}u^\lambda_n = f.$$

By Itô's formula, for any stopping times $\tau_1\leqslant\tau_2$,

$$\begin{aligned}&u^\lambda_n(X^n_{\tau_2}) - u^\lambda_n(X^n_{\tau_1})\\ =&\int_{\tau_1}^{\tau_2}\int_{\mathbb{R}^d}\int_0^{\kappa_n(X_{s-},z)}[u^\lambda_n(X^n_{s-}+z) - u^\lambda_n(X^n_{s-})]\widetilde{N}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\\ &+\lambda\int_{\tau_1}^{\tau_2}u^\lambda_n(X^n_s)\mathrm{d}s - \int_{\tau_1}^{\tau_2}f(X^n_s)\mathrm{d}s.\end{aligned}$$

Hence,

$$\int_{\tau_1}^{\tau_2} f(X_s^n)\mathrm{d}s = \int_{\tau_1}^{\tau_2}\int_{\mathbb{R}^d}\int_0^{\kappa_n(X_{s-}^n,z)} [u_n^\lambda(X_{s-}^n+z)-u_n^\lambda(X_{s-}^n)]\widetilde{N}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) + u_n^\lambda(X_{\tau_1}^n)-u_n^\lambda(X_{\tau_2}^n)+\lambda\int_{\tau_1}^{\tau_2} u_n^\lambda(X_s^n)\mathrm{d}s.$$

Denote

$$M_t^n := \int_0^t\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{\{r<\kappa_n(X_{s-}^n,z)\}}[u_n^\lambda(X_{s-}^n+z)-u_n^\lambda(X_{s-}^n)]\widetilde{N}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s).$$

By Burkholder-Davis-Gundy's inequality, we get that for any $\delta>0$, $m\in\mathbb{N}_+$ and bounded stopping time τ,

$$\begin{aligned}
&\mathbf{E}\left|\int_\tau^{\tau+\delta} f(X_s^n)\mathrm{d}s\right|^2\\
\leqslant& C_m\left\{\mathbf{E}([M^n]_{\tau+\delta}-[M^n]_\tau)+\|u_n^\lambda\|_\infty^2+(\lambda\delta\|u_n^\lambda\|_\infty)^2\right\}\\
\leqslant& C_m\left\{\mathbf{E}([M^n]_{\tau+\delta}-[M^n]_\tau)+[1+(\lambda\delta)^2]\|u_n^\lambda\|_\infty^2\right\}.
\end{aligned}\tag{6.47}$$

On the other hand,

$$\begin{aligned}
&[M^n]_{\tau+\delta}-[M^n]_\tau\\
=&\int_\tau^{\tau+\delta}\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{\{r<\kappa_n(X_{s-}^n,z)\}}[u_n^\lambda(X_{s-}^n+z)-u_n^\lambda(X_{s-}^n)]^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\\
\leqslant& C\int_\tau^{\tau+\delta}\int_{\mathbb{R}^d}\int_0^{\Lambda_2}(|z|^2\|\nabla u_n^\lambda\|_\infty^2\wedge\|u_n^\lambda\|_\infty^2)N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\\
=&C\int_\tau^{\tau+\delta}\int_{\mathbb{R}^d}\int_0^{\Lambda_2} g_n^\lambda(z)\widetilde{N}(\mathrm{d}r,\mathrm{d}z;\mathrm{d}s)+C\int_\tau^{\tau+\delta}\int_{\mathbb{R}^d}\int_0^{\Lambda_2} g_n^\lambda(z)\mathrm{d}r\frac{\mathrm{d}z}{|z|^{d+\alpha}}\mathrm{d}s,
\end{aligned}\tag{6.48}$$

where

$$g_n^\lambda(z) := |z|^2\|\nabla u_n^\lambda\|_\infty^2\wedge\|u_n^\lambda\|_\infty^2.$$

By Theorem 6.1 and interpolation, we have

$$\|u_n^\lambda\|_\infty\lesssim\lambda^{-\theta}\|f\|_{\mathscr{C}^\mu},\quad \|\nabla u_n^\lambda\|_\infty\lesssim\lambda^{\frac{1}{\alpha}-\theta}\|f\|_{\mathscr{C}^\mu},\quad\forall\mu\in(-(\tfrac{\alpha-1}{2}\wedge\vartheta),\beta],\theta\in(0,1+\tfrac{\mu}{\alpha}).\tag{6.49}$$

This yields

$$|g_n^\lambda(z)|\lesssim\|f\|_{\mathscr{C}^\mu}^2(|z|^2\lambda^{-2\theta+\frac{2}{\alpha}}\wedge\lambda^{-2\theta}),\quad\forall\mu\in(-(\tfrac{\alpha-1}{2}\wedge\vartheta),\beta],\theta\in(0,1+\tfrac{\mu}{\alpha}).\tag{6.50}$$

For any $\delta \leqslant \lambda_0^{-1}$, choosing $\lambda = \delta^{-1}$ and combining (6.47)-(6.50), we get

$$\begin{aligned} &\mathbf{E}\left|\int_\tau^{\tau+\delta} f(X_s^n)\mathrm{d}s\right|^2 \\ \leqslant& C\delta\int_{\mathbb{R}^d} g_n^\lambda(z)\frac{\mathrm{d}z}{|z|^{d+\alpha}} + C\|u_n^\lambda\|_\infty^2 \\ \leqslant& C\|f\|_{\mathscr{C}^\mu}^2\left(\delta\lambda^{-2\theta+\frac{2}{\alpha}}\int_{|z|<\lambda^{-1/\alpha}}|z|^{2-d-\alpha}\mathrm{d}z + \delta\lambda^{-2\theta}\int_{|z|\geqslant\lambda^{-1/\alpha}}|z|^{-d-\alpha}\mathrm{d}zC + \lambda^{-2\theta}\right) \\ \leqslant& C\|f\|_{\mathscr{C}^\mu}^2\delta\lambda^{1-2\theta} = C\|f\|_{\mathscr{C}^\mu}^2\delta^{2\theta}, \end{aligned} \tag{6.51}$$

here C is independent with n. Let $A_t^n := \int_0^t b_n(X_s^n(x))\mathrm{d}s$ and $\mathcal{T}$ be the collection of all bounded stopping time. The above estimate and Burkholder-Davis-Gundy's inequality yield

$$\begin{aligned} &\sup_{\tau\in\mathcal{T}}\mathbf{E}|X_{\tau+\delta}^n - X_\tau^n| \\ \leqslant& \sup_{\tau\in\mathcal{T}}\mathbf{E}\left(|A_{\tau+\delta}^n - A_\tau^n| + \left|\int_\tau^{\tau+\delta}\int_{\mathbb{R}^d}\int_0^{\kappa_n(X_{s-}^n,z)} z\widetilde{N}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right|\right) \\ \leqslant& \sup_{\tau\in\mathcal{T}}\left(\mathbf{E}\left|\int_\tau^{\tau+\delta} b_n(X_s^n)\mathrm{d}s\right|^2\right)^{\frac{1}{2}} + C\sup_{\tau\in\mathcal{T}}\mathbf{E}\left(\int_\tau^{\tau+\delta}\int_{\mathbb{R}^d}\int_0^{\Lambda_2}|z|^2N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right)^{\frac{1}{2}} \\ \leqslant& C\|b\|_{\mathscr{C}^\beta}\delta^\theta + C\sup_{\tau\in\mathcal{T}}\mathbf{E}\left[\int_{|z|\leqslant 1}|z|^2N_\tau^\delta(\mathrm{d}z)\right]^{\frac{1}{2}} + C\sup_{\tau\in\mathcal{T}}\mathbf{E}\left[\int_{|z|>1}|z|^2N_\tau^\delta(\mathrm{d}z)\right]^{\frac{1}{2}}. \end{aligned}$$

where

$$N_\tau^\delta(\mathrm{d}z) := \int_\tau^{\tau+\delta}\int_0^{\Lambda_2} N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s),$$

it is not hard to see that N_τ^δ is a Poisson random measure on $\mathbb{R}^d$ with intensity measure $\delta\Lambda_2\frac{\mathrm{d}z}{|z|^{d+\alpha}}$. Notice that for fixed $\omega\in\Omega$, N_τ^δ is a counting measure, by the elementary inequality: $(\sum_k |a_k|^p)^{1/p} \leqslant (\sum_k |a_k|^q)^{1/q}$, $\forall\, p \geqslant q > 0$ and $\{a_k\}\subset\mathbb{R}$, we also have

$$\left(\int_{|z|>1}|z|^2N_\tau^\delta(\mathrm{d}z)\right)^{\frac{1}{2}} \leqslant \int_{|z|>1}|z|N_\tau^\delta(\mathrm{d}z).$$

Thus, for small $\delta \leqslant \lambda_0^{-1}$ we have

$$\begin{aligned} &\sup_{\tau\in\mathcal{T}}\mathbf{E}|X_{\tau+\delta}^n - X_\tau^n| \\ \leqslant& C\|b\|_{\mathscr{C}^\beta}\delta^\theta + C\sup_{\tau\in\mathcal{T}}\left[\mathbf{E}\int_{|z|\leqslant 1}|z|^2N_\tau^\delta(\mathrm{d}z)\right]^{\frac{1}{2}} + C\sup_{\tau\in\mathcal{T}}\mathbf{E}\int_{|z|>1}|z|N_\tau^\delta(\mathrm{d}z) \\ \leqslant& C(\|b\|_{\mathscr{C}^\beta}\delta^\theta + \delta^{\frac{1}{2}} + \delta) \lesssim \delta^{\frac{1}{2}}, \end{aligned}$$

and consequently

$$\lim_{\delta\downarrow 0}\mathbf{E}\sup_{\tau\in\mathcal{T}}|X^n_{\tau+\delta}-X^n_\tau|=0.$$

By Aldous tightness criterion, we obtain that $\{\mathbb{P}^n_x:=\mathbf{P}\circ(X^n_t)^{-1}\}_{n\in\mathbb{N}}$ is tight. So, upon taking a subsequence, still denote by n, we can assume that $\mathbb{P}^n_x\Rightarrow\mathbb{P}_x$. By (6.51), we also have

$$\mathbb{E}_x\left|\int_{t_0}^{t_1}f(\omega_s)\mathrm{d}s\right|^2=\lim_{n\to\infty}\mathbb{E}^n_x\left|\int_{t_0}^{t_1}f(\omega_s)\mathrm{d}s\right|^2\leqslant C\|f\|^2_{\mathscr{C}^\mu}|t_1-t_0|^{2\theta},$$

where $\mu\in(-(\frac{\alpha-1}{2}\wedge\vartheta),\beta]$ and $\theta\in(0,1+\frac{\mu}{\alpha})$, i.e. $\mathbb{P}_x\in\mathscr{K}^\mu(\mathbb{D})$. Hence, by Proposition 6.21, for any $f\in C_b^\infty$, we can define

$$A_t^{b\cdot\nabla f}(\omega):=\lim_{n\to\infty}\int_0^t b_n\cdot\nabla f(\omega_s)\mathrm{d}s,\quad \mathbb{P}_x-a.s..$$

Next we verify that $\mathbb{P}_x\in\mathscr{M}^\mu_{\kappa,b}(x)$ with $\mu\in(-(\frac{\alpha-1}{2}\wedge\vartheta),\beta)$. Let $\mathcal{B}^0_t:=\sigma(\{\omega_s:\omega\in\mathbb{D},s\leqslant t\})$, $\mathcal{B}_t=\cap_{s>t}\mathcal{B}^0_s$, $\mathcal{B}=\sigma(\cup_{t\in\mathbb{R}_+}\mathcal{B}_t)$, $\mathrm{D}_{\mathbb{P}_x}:=\{t>0:\mathbb{P}_x(\omega_t=\omega_{t-})<1\}$. For any $s,s_i,t\in\mathrm{D}_{\mathbb{P}_x}$, $0\leqslant s_1\leqslant s_2\leqslant\cdots\leqslant s_k\leqslant s\leqslant t$, $f\in C_b^\infty$ and $h_1,h_2,\cdots,h_k\in C_b(\mathbb{R}^d)$, denote $H:=\Pi^k_{i=1}h_i(\omega_{s_i})\in\mathcal{B}_s$, then

$$\begin{aligned}
&|\mathbb{E}_x[(M^f_t-M^f_s)\Pi^k_{i=1}h_i(\omega_{s_i})]\\
\leqslant&\left|(\mathbb{E}_x-\mathbb{E}^n_x)\left[f(\omega_t)-f(\omega_s)-\int_s^t(\mathscr{L}^\alpha_{\kappa_m}f+b_m\cdot\nabla f)(\omega_r)\mathrm{d}r\right]H\right|\\
&+\left|\mathbb{E}_x\left[\int_s^t(\mathscr{L}^\alpha_{\kappa_m}-\mathscr{L}^\alpha_\kappa)f(\omega_r)\mathrm{d}r+\int_s^t b_m\cdot\nabla f(\omega_r)\mathrm{d}r-(A_t^{b\cdot\nabla f}-A_s^{b\cdot\nabla f})\right]H\right|\\
&+\left|\mathbb{E}^n_x\left[f(\omega_t)-f(\omega_s)-\int_s^t(\mathscr{L}^\alpha_{\kappa_n}f+b_n\cdot\nabla f)(\omega_r)\mathrm{d}r\right]H\right|\\
&+\left|\mathbb{E}^n_x\left[\int_s^t[(\mathscr{L}^\alpha_{\kappa_n}-\mathscr{L}^\alpha_{\kappa_m})f+(b_n-b_m)\cdot\nabla f](\omega_r)\mathrm{d}r\right]H\right|.
\end{aligned}\tag{6.52}$$

Notice that for any m, the first term on the right side of (6.52) goes to 0 as n goes to 0. Since $\mathbb{P}_x\in\mathscr{K}^\mu(\mathbb{D})$, by the definition of $A^{b\cdot\nabla f}_t$, we have

$$\begin{aligned}
&\lim_{m\to\infty}\left|\mathbb{E}_x\left[\int_s^t(\mathscr{L}^\alpha_{\kappa_m}-\mathscr{L}^\alpha_\kappa)f(\omega_r)\mathrm{d}r+\int_s^t b_m\cdot\nabla f(\omega_r)\mathrm{d}r-(A_t^{b\cdot\nabla f}-A_s^{b\cdot\nabla f})\right]H\right|\\
\leqslant&\Pi^k_{i=1}\|h_i\|_\infty\lim_{m\to\infty}\left(\left|\mathbb{E}_x\int_s^t(\mathscr{L}^\alpha_{\kappa_m}-\mathscr{L}^\alpha_\kappa)f(\omega_r)\mathrm{d}r\right|\right.\\
&\left.+\mathbb{E}_x\left|\int_s^t b_m\cdot\nabla f(\omega_r)\mathrm{d}r-(A_t^{b\cdot\nabla f}-A_s^{b\cdot\nabla f})\right|\right)\\
=&0.
\end{aligned}$$

Similarly, the fourth term goes to 0 uniformly in n as m goes to 0. And by definition, the third term on the right side of (6.52) is zero. Thus, letting first $n\to\infty$ and then $m\to\infty$ on the right side of (6.52), we get

$$\mathbb{E}_x[(M^f_t-M^f_s)\Pi^k_{i=1}h_i(\omega_{s_i})]=0,\quad\forall s,s_i,t\in\mathrm{D}_{\mathbb{P}_x},s_i\leqslant s\leqslant t.$$

By [20, Lemma 7.7 of Chapter 3], $\mathrm{D}_{\mathbb{P}_x}$ is at most countable, noticing that M_t^f is càdlàg under $\mathbb{P}_x$, we obtain

$$\mathbb{E}_x[(M_t^f - M_s^f)\Pi_{i=1}^k h_i(\omega_{s_i})] = 0, \quad \forall s, s_i, t \in [0,\infty), s_i \leqslant s \leqslant t.$$

□

We close this section by giving the definition of weak solution.

Definition 6.25 (Weak solution)**.** *Let $\beta \in \mathbb{R}$, $\alpha \in (0,2)$. We say that $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P}, X, N, A)$ is a weak solution to*

$$X_t = x + \int_0^t \int_{\mathbb{R}^d} \int_0^{\kappa(X_{s-},z)} z N^{(\alpha)}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) + \int_0^t b(X_s)\mathrm{d}s, \tag{6.53}$$

if

1. *$(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ is a complete filtered probability space and X_t, A_t are càdlàg processes adapted with $\mathcal{F}_t$. N is a Poisson random measure and for any compact set $B \subseteq \mathbb{R}_+ \times \mathbb{R}^d \backslash \{0\}$, $N(B;t)$ is a $\mathcal{F}_t$ adapted Poisson process with intensity $\int_{\mathbb{R}^d} \int_0^\infty \mathbf{1}_B(r,z)\mathrm{d}r \frac{\mathrm{d}z}{|z|^{d+\alpha}}$;*
2. $$X_t = x + \int_0^t \int_{\mathbb{R}^d} \int_0^\infty z \mathbf{1}_{[0,\kappa(X_{s-},z))} N^{(\alpha)}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) + A_t,$$
   and for any $b_n \in C_b^\infty$ and $b_n \xrightarrow{\mathscr{C}^\beta} b$, we have
   $$\int_0^t b_n(X_s)\mathrm{d}s \longrightarrow A_t$$
   in probability $\mathbf{P}$ uniformly over bounded time intervals;
3. *there are constant $\gamma, C > 0$ such that*
   $$\mathbf{E}\,|A_t - A_s|^2 \leqslant C|t-s|^{1+\gamma}, \quad s,t \in [0,T].$$

Thanks to the martingale representation theorem for Poisson noise(see II.1.c on p.74 of [42]), following the argument in [42, Theorem II_{10}] and [76, Proposition 3.13], we have the equivalence between martingale solution and weak solution without any surprise. In order to make thesis complete, we show the equivalence in Appendix.

Theorem 6.26. *Let $\mathbb{P} \in \mathscr{P}(\mathbb{D})$,*

1. *if $\alpha \in (0,1]$, $b \in \mathscr{C}^\beta$ with $\beta > 0$, then $\mathbb{P} \in \mathscr{M}_{\kappa,b}(x)$ if and only if there is a weak solution $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P}, X, N, A)$ so that $\mathbf{P} \circ X^{-1} = \mathbb{P}$;*
2. *if $\alpha \in (1,2)$, $b \in \mathscr{C}^\beta$ with $\beta \leqslant 0$, then $\mathbb{P} \in \mathscr{M}^\mu_{\kappa,b}(x)$ for some $\mu < \beta$ if and only if there is a weak solution $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P}, X, N, A)$ so that $\mathbf{P} \circ X^{-1} = \mathbb{P} \in \mathscr{K}^\mu(\mathbb{D})$.*

6.5 Regularity of densities of weak solutions

Thanks to Theorem 6.26, it is equivalent to consider the weak solution of (6.53) and martingale solution associated with $\mathscr{L}^\alpha_{\kappa,b}$. We are going to prove that the law of the weak solution of (6.53) has a density in some Besov space under some mild assumptions. Most results in this section are inspired by Debussche and Fournier's work [15].

Through out this section, we assume ν satisfies the following assumption for some $\alpha \in (0,2)$:

Assumption 7. *(i)* $\int_{|z|\geqslant 1} |z|^p \nu(\mathrm{d}z) < \infty, \quad \forall p \in [0,\alpha)$,

(ii) there exists $C > 0$ *such that* $\int_{|z|\leqslant a} |z|^2 \nu(\mathrm{d}z) \leqslant C a^{2-\alpha}, \quad \forall a \in (0,1]$,

(iii) there exists $c > 0$ *such that* $\int_{|z|\leqslant a} |\langle \xi, z\rangle|^2 \nu(\mathrm{d}z) \geqslant c a^{2-\alpha}, \ \forall \xi \in \mathbb{S}^{d-1}, a \in (0,1]$.

Define

$$N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) := \begin{cases} N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) & \alpha \in (0,1) \\ N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) - \mathrm{d}r\mathbf{1}_{B_1}(z)\nu(\mathrm{d}z)\mathrm{d}s & \alpha = 1 \\ N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) - \mathrm{d}r\nu(\mathrm{d}z)\mathrm{d}s & \alpha \in (1,2), \end{cases}$$

where N is a a Poisson random measure on $\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}_+$ with intensity measure $\mathrm{d}r\nu(\mathrm{d}z)\mathrm{d}s$. We also assume Y_t solves the following equation:

$$Y_t = Y_0 + \int_0^t a(Y_s)\mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} \int_0^\infty g(Y_{s-},z)\mathbf{1}_{[0,k(Y_{s-},z)]}(r) N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s), \tag{6.54}$$

where a, g, k are bounded measurable functions.

Lemma 6.27. *Assume*

$$|a(y)| \leqslant c_0, |g(y,z)| \leqslant c_2|z|, |k(y,z)| \leqslant \lambda_2$$

and Y solves (6.54). Then for all $p \in (0,\alpha)$ and $0 \leqslant s \leqslant t \leqslant s+1$ we have

$$\mathbf{E} \sup_{v\in[s,t]} |Y_v - Y_s|^p + \mathbf{E} \sup_{v\in[s,t]} |Y_{v-} - Y_s|^p \leqslant C(p,c_0,c_2,\lambda_2)|t-s|^{\frac{p}{\alpha\vee 1}}. \tag{6.55}$$

Furthermore, if $\alpha \in (0,1)$ and $p \in [\alpha,1)$, then for all $0 \leqslant s \leqslant t \leqslant s+1$ we have

$$\mathbf{E}\left(\sup_{v\in[s,t]} |Y_v - Y_s|^p \wedge 1\right) \leqslant C(p,c_0,c_2,\lambda_2)|t-s|^p. \tag{6.56}$$

Proof. For all $0 < p < \alpha$ and $0 \leqslant s \leqslant t \leqslant s+1$, we have: if $\alpha \in (1,2)$

$$\mathbf{E}\left|\int_s^t a(Y_u)\mathrm{d}u\right|^p \leqslant C|t-s|^p \leqslant C|t-s|^{\frac{p}{\alpha}}, \tag{6.57}$$

and if $\alpha \in (0,1]$

$$\mathbf{E}\left|\int_s^t a(Y_u)\mathrm{d}u\right|^p \leqslant C|t-s|^p. \tag{6.58}$$

Then the inequality (6.55) is a simple consequence of (6.57) and (6.58) and the following inequality:

$$\begin{aligned}&\mathbf{E}\left[\sup_{v\in[s,t]}\left|\int_s^v\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(Y_{u-},z)]}(r)g(Y_{u-},z)N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}u)\right|^p\right]\\ \leqslant & C(p,c_2,\lambda_2)|t-s|^{p/\alpha},\end{aligned} \tag{6.59}$$

for all $p\in(0,\alpha)$ and $0\leqslant s\leqslant t\leqslant s+1$. Actually, if $\alpha\in(1,2)$, write

$$\begin{aligned}&\int_s^v\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(Y_{u-},z)]}(r)g(Y_{u-},z)N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z;\mathrm{d}u)\\ =&I_1(v)+I_2(v):=\int_s^v\int_{|z|\leqslant|t-s|^{1/\alpha}}\int_0^{k(Y_{u-},z)} g(Y_{u-},z)N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}u)\\ &+\int_s^v\int_{|z|>|t-s|^{1/\alpha}}\int_0^{k(Y_{u-},z)} g(Y_{u-},z)N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}u),\end{aligned}$$

For I_1, notice that $\frac{p}{2}<1$, by Burkholder-Davis-Gundy's inequality,

$$\begin{aligned}\mathbf{E}\left[\sup_{v\in[s,t]}|I_1(v)|^p\right]\leqslant& C_p\mathbf{E}\left[\left|\int_s^t\int_{|z|\leqslant|t-s|^{1/\alpha}}\int_0^{k(Y_{u-},z)}|g(Y_{u-},z)|^2N(\mathrm{d}r,\mathrm{d}z;\mathrm{d}u)\right|^{\frac{p}{2}}\right]\\ \leqslant& C_p\left[\mathbf{E}\int_s^t\int_{|z|\leqslant|t-s|^{1/\alpha}}\int_0^{k(Y_{u-},z)}|g(Y_{u-},z)|^2N(\mathrm{d}r,\mathrm{d}z;\mathrm{d}u)\right]^{\frac{p}{2}}\\ \leqslant& C_p\lambda_2\left[c_2^2|t-s|\int_{|z|\leqslant|t-s|^{1/\alpha}}|z|^2\nu(\mathrm{d}z)\right]^{\frac{p}{2}}\leqslant C(p,c_2,\lambda_2)|t-s|^{p/\alpha}.\end{aligned} \tag{6.60}$$

For I_2, similarly, we have

$$\begin{aligned}c_2\mathbf{E}\left[\sup_{v\in[s,t]}|I_2(v)|^p\right]\leqslant& C_p\mathbf{E}\left[\left|\int_s^t\int_{|z|>|t-s|^{1/\alpha}}\int_0^{k(Y_{u-},z)}|g(Y_{u-},z)|^2N(\mathrm{d}r,\mathrm{d}z;\mathrm{d}u)\right|^{\frac{p}{2}}\right]\\ \leqslant& C_pc_2^2\mathbf{E}\left[\left|\int_s^t\int_{|z|>|t-s|^{1/\alpha}}\int_0^{\lambda_2}|z|^2N(\mathrm{d}r,\mathrm{d}z;\mathrm{d}u)\right|^{\frac{p}{2}}\right].\end{aligned}$$

Let $N_{s,t}(\mathrm{d}z)=\int_s^t\int_0^{\lambda_2}N(\mathrm{d}r,\mathrm{d}z;\mathrm{d}u)$, then $N_{s,t}$ is a Poisson random measure with intensity $\lambda_2|t-s|\nu(\mathrm{d}z)$. Notice that $N_{s,t}$ is a counting measure, by the elementary inequality:

$(\sum_k |a_k|^p)^{1/p} \leqslant (\sum_k |a_k|^q)^{1/q}$, $\forall\, p \geqslant q > 0, \{a_k\} \subset \mathbb{R}$ and Lemma A.1 of [15], we obtain

$$\begin{aligned}\mathbf{E}\left[\sup_{v\in[s,t]} |I_2(v)|^p\right] \leqslant & C_p c_2^p \mathbf{E}\left[\left|\int_{|z|>|t-s|^{1/\alpha}} |z|^2 N_{s,t}(\mathrm{d}z)\right|^{\frac{p}{2}}\right] \\ \leqslant & C_p c_2^p \mathbf{E}\int_{|z|>|t-s|^{1/\alpha}} |z|^p N_{s,t}(\mathrm{d}z) \\ \leqslant & C_p c_2^p \lambda_2 |t-s| \int_{|z|>|t-s|^{1/\alpha}} |z|^p \nu(\mathrm{d}z) \leqslant C(p, c_2, \lambda_2)|t-s|^{p/\alpha}.\end{aligned} \tag{6.61}$$

Combining (6.60) and (6.61), we get the desired result for $\alpha \in (1,2)$. By the similar argument we get that for $0 < p < \alpha \leqslant 1$ and $0 \leqslant s \leqslant t \leqslant s+1$

$$\begin{aligned}&\mathbf{E}\left[\sup_{v\in[s,t]}\left|\int_s^v\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(Y_{u-},z)]}(r) g(Y_{u-},z) N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}u)\right|^p\right] \\ \leqslant & C(p,c_2,\lambda_2)|t-s|^{p/\alpha}.\end{aligned}$$

Now we only need to show that for $p \in [\alpha, 1)$ and $0 \leqslant s \leqslant t \leqslant s+1$, (6.56) holds. Since $\alpha p < p$, we have

$$\begin{aligned}&|Y_t - Y_s|^p \wedge 1 \\ \leqslant & \left|\int_s^t a(Y_u) du\right|^p + \left|\int_s^t\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(Y_{u-},z)]}(r) g(Y_{u-},z) N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}u)\right|^p \wedge 1 \\ \leqslant & C|t-s|^p + \left|\int_s^t\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(Y_{u-},z)]}(r) g(Y_{u-},z) N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}u)\right|^{\alpha p}.\end{aligned} \tag{6.62}$$

By (6.59), we get (6.56). □

Lemma 6.28. *Suppose $\theta_i \in (0,1), i = 1,2,3$ and $c_j > 0, j = 0,1,2,3$,*

$$\begin{aligned}|a(y)| \leqslant c_0, \quad & |a(y_1) - a(y_2)| \leqslant c_1 |y_1 - y_2|^{\theta_1}, \\ |g(y,z)| \leqslant c_2|z|, \quad & |g(y_1,z) - g(y_2,z)| \leqslant c_3 |y_1-y_2|^{\theta_2}|z|,\end{aligned} \tag{6.63}$$

k satisfies ($\mathbf{H}_2$), ($\mathbf{H}_3$) with Λ_i and ϑ replaced by λ_i and θ_3, respectively. For any $\epsilon \in (0, t\wedge 1)$, we can find a $\mathcal{F}_{t-\epsilon}$-measurable variable V_t^ϵ such that for all $p \in (0,\alpha)$

$$\mathbf{E}|Y_t - Y_t^\epsilon|^p \leqslant C\epsilon^{\theta_0 p}, \tag{6.64}$$

where

$$Y_t^\epsilon = V_t^\epsilon + \int_{t-\epsilon}^t\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(Y_{t-\epsilon},z)]}(r) g(Y_{t-\epsilon},z) N^\alpha(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s),$$

and if $\alpha \in [1,2)$,

$$\theta_0 = \frac{1}{\alpha}\big[(\alpha+\theta_1)\wedge(1+\theta_2)\wedge(1+\tfrac{\theta_3}{\alpha})\big],$$

if $\alpha \in (0,1)$

$$\theta_0 = \frac{1}{1-\theta_1} \wedge \frac{1}{\alpha}\big[(\alpha+\theta_1)\wedge(1+\theta_2)\wedge(1+\theta_3)\big],$$

Proof. We first prove the case when $\alpha \in (1,2)$. Take

$$V_t^\epsilon := Y_{t-\epsilon} + \epsilon a(Y_{t-\epsilon}),$$

then

$$\begin{aligned} Y_t - Y_t^\epsilon =& I_{t,\epsilon} + J_{t,\epsilon} := \int_{t-\epsilon}^t [a(Y_s) - a(Y_{t-\epsilon})]\mathrm{d}s \\ &+ \int_{t-\epsilon}^t \int_{\mathbb{R}^d} \int_0^\infty \Big[g(Y_{s-},z)\mathbf{1}_{[0,k(Y_{s-},z)]}(r) - g(Y_{t-\epsilon},z)\mathbf{1}_{[0,k(Y_{t-\epsilon},z)]}(r) \Big] N^\alpha(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s). \end{aligned}$$

For all $p \in (0,1]$, by Jensen's inequality,

$$\begin{aligned} \mathbf{E}[|I_{t,\epsilon}|^p] &\leqslant \|a\|_{C^{\theta_1}}^p \mathbf{E} \left(\int_{t-\epsilon}^t |Y_s - Y_{t-\epsilon}|^{\theta_1} \mathrm{d}s \right)^p \\ &\leqslant \|a\|_{C^{\theta_1}}^p \left(\int_{t-\epsilon}^t \mathbf{E}|Y_s - Y_{t-\epsilon}|^{\theta_1} \mathrm{d}s \right)^p \overset{(6.55)}{\leqslant} C\epsilon^{p(1+\frac{\theta_1}{\alpha})}. \end{aligned}$$

If $p \in (1,\alpha)$, by Hölder's inequality,

$$\begin{aligned} \mathbf{E}[|I_{t,\epsilon}|^p] &\leqslant \|a\|_{C^{\theta_1}}^p \mathbf{E} \left[\left| \int_{t-\epsilon}^t |Y_s - Y_{t-\epsilon}|^{\theta_1} \mathrm{d}s \right|^p \right] \\ &\leqslant \|a\|_{C^{\theta_1}}^p \epsilon^{p-1} \mathbf{E} \int_{t-\epsilon}^t |Y_s - Y_{t-\epsilon}|^{p\theta_1} \mathrm{d}s \overset{(6.55)}{\leqslant} C\epsilon^{p(1+\frac{\theta_1}{\alpha})}. \end{aligned}$$

To sum up, for each $p \in (0,\alpha)$,

$$\mathbf{E}[|I_{t,\epsilon}|^p] \leqslant C\epsilon^{p(1+\frac{\theta_1}{\alpha})}. \tag{6.65}$$

For $J_{t,\epsilon}$,

$$\begin{aligned} J_{t,\epsilon} =& \int_{t-\epsilon}^t \int_{\mathbb{R}^d} \int_0^{k(Y_{t-\epsilon},z)} \big[g(Y_{s-},z) - g(Y_{t-\epsilon},z) \big] N^\alpha(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) \\ &+ \int_{t-\epsilon}^t \int_{\mathbb{R}^d} \int_{k(Y_{t-\epsilon},z)}^{k(Y_{s-},z)} g(Y_{s-},z) N^\alpha(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) =: J_{t,\epsilon}^1 + J_{t,\epsilon}^2, \end{aligned} \tag{6.66}$$

where we abuse the notation $\int_u^v = -\int_v^u$ when $u > v$. Notice that $p \in (0,\alpha)$, like the proof of Lemma 6.27, one can see that

$$\begin{aligned}
\mathbf{E}[|J^1_{t,\epsilon}|^p] \leqslant& C\mathbf{E}\left[\left|\int_{t-\epsilon}^t\int_{\mathbb{R}^d}\int_0^{\lambda_2}\left|g(Y_{s-},z)-g(Y_{t-\epsilon},z)\right|^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right|^{\frac{p}{2}}\right]\\
\leqslant& C\mathbf{E}\left[\left|\int_{t-\epsilon}^t\int_{\mathbb{R}^d}\int_0^{\lambda_2}\left|Y_{s-}-Y_{t-\epsilon}\right|^{2\theta_2}|z|^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right|^{\frac{p}{2}}\right]\\
\leqslant& C\mathbf{E}\left[\left|\int_{t-\epsilon}^t\int_{|z|\leqslant\epsilon^{\frac{1}{\alpha}}}\int_0^{\lambda_2}\left|Y_{s-}-Y_{t-\epsilon}\right|^{2\theta_2}|z|^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right|^{\frac{p}{2}}\right]\\
&+C\mathbf{E}\left[\left|\int_{t-\epsilon}^t\int_{|z|>\epsilon^{\frac{1}{\alpha}}}\int_0^{\lambda_2}\left|Y_{s-}-Y_{t-\epsilon}\right|^{2\theta_2}|z|^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right|^{\frac{p}{2}}\right]\\
\leqslant& C\mathbf{E}\left[\sup_{s\in[t-\epsilon,t]}|Y_{s-}-Y_{t-\epsilon}|^{p\theta_2}\left(\int_{t-\epsilon}^t\int_{|z|\leqslant\epsilon^{\frac{1}{\alpha}}}\int_0^{\lambda_2}|z|^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right)^{\frac{p}{2}}\right]\\
&+C\mathbf{E}\left(\int_{t-\epsilon}^t\int_{|z|>\epsilon^{\frac{1}{\alpha}}}\int_0^{\lambda_2}\left|Y_{s-}-Y_{t-\epsilon}\right|^{p\theta_2}|z|^p N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right)\\
\leqslant& C\left[\mathbf{E}\sup_{s\in[t-\epsilon,t]}|Y_{s-}-Y_{t-\epsilon}|^{\alpha\theta_2}\right]^{\frac{p}{\alpha}}\left[\mathbf{E}\left(\int_{t-\epsilon}^t\int_{|z|\leqslant\epsilon^{\frac{1}{\alpha}}}\int_0^{\lambda_2}|z|^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right)^{\frac{\alpha p}{2(\alpha-p)}}\right]^{1-\frac{p}{\alpha}}\\
&+C\int_{t-\epsilon}^t\int_{|z|>\epsilon^{\frac{1}{\alpha}}}\int_0^{\lambda_2}\mathbf{E}|Y_{s-}-Y_{t-\epsilon}|^{p\theta_2}|z|^p\mathrm{d}r\nu(\mathrm{d}z)\mathrm{d}u\overset{(6.55)}{\leqslant}C\epsilon^{\frac{p}{\alpha}(1+\theta_2)}.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\mathbf{E}[|J^2_{t,\epsilon}|^p] \leqslant& C\mathbf{E}\left[\left|\int_{t-\epsilon}^t\int_{\mathbb{R}^d}\int_{k(Y_{t-\epsilon},z)}^{k(Y_{s-},z)}|z|^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right|^{\frac{p}{2}}\right]\\
\leqslant& C\left|\mathbf{E}\int_{t-\epsilon}^t\int_{|z|\leqslant\epsilon^{\frac{1}{\alpha}(1+\frac{\theta_3}{\alpha})}}\int_{k(Y_{t-\epsilon},z)}^{k(Y_{s-},z)}|z|^2 N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right|^{\frac{p}{2}}\\
&+C\mathbf{E}\int_{t-\epsilon}^t\int_{|z|>\epsilon^{\frac{1}{\alpha}(1+\frac{\theta_3}{\alpha})}}\int_{k(Y_{t-\epsilon},z)}^{k(Y_{s-},z)}|z|^p N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\\
\leqslant& C\left[\int_{t-\epsilon}^t\int_{|z|\leqslant\epsilon^{\frac{1}{\alpha}(1+\frac{\theta_3}{\alpha})}}\mathbf{E}|k(Y_{s-},z)-k(Y_{t-\epsilon},z)||z|^2\nu(\mathrm{d}z)\mathrm{d}s\right]^{\frac{p}{2}}\\
&+C\int_{t-\epsilon}^t\int_{|z|>\epsilon^{\frac{1}{\alpha}(1+\frac{\theta_3}{\alpha})}}\mathbf{E}|k(Y_{s-},z)-k(Y_{t-\epsilon},z)||z|^p\nu(\mathrm{d}z)\mathrm{d}s\\
\leqslant& C\left[\int_{t-\epsilon}^t\int_{|z|\leqslant\epsilon^{\frac{1}{\alpha}(1+\frac{\theta_3}{\alpha})}}\mathbf{E}|Y_{s-}-Y_{t-\epsilon}|^{\theta_3}|z|^2\nu(\mathrm{d}z)\mathrm{d}s\right]^{\frac{p}{2}}\\
&+C\int_{t-\epsilon}^t\int_{|z|>\epsilon^{\frac{1}{\alpha}(1+\frac{\theta_3}{\alpha})}}\mathbf{E}|Y_{s-}-Y_{t-\epsilon}|^{\theta_3}|z|^p\nu(\mathrm{d}z)\mathrm{d}s\overset{(6.55)}{\leqslant}C\epsilon^{\frac{p}{\alpha}(1+\frac{\theta_3}{\alpha})}.
\end{aligned}$$

Combing the above inequalities, we get

$$\mathbf{E}[|J_{t,\epsilon}|^p] \leqslant C\epsilon^{\frac{p}{\alpha}(1+\theta_2\wedge\frac{\theta_3}{\alpha})}. \tag{6.67}$$

Thus we get (6.64) for $\alpha \in (1,2)$.

For $\alpha \in (0,1)$, let $\delta = \epsilon^{1/(1-\theta_1)}$, $s \in [t-\epsilon, t]$, $s_\delta = t-\epsilon+\delta\lfloor(s-(t-\epsilon))/\delta\rfloor$, here $\lfloor a\rfloor$ is the max integer less than or equal to a. Consider the solution to

$$V_u^\epsilon = Y_{t-\epsilon} + \int_{t-\epsilon}^u b(V_{s_\delta}^\epsilon)\mathrm{d}s, \quad u \in [t-\epsilon, t].$$

One can see that V_t^ϵ is well defined and $\mathcal{F}_{t-\epsilon}$ measurable. Writing

$$V_u^\epsilon = Y_{t-\epsilon} + \int_{t-\epsilon}^u b(V_s^\epsilon)\mathrm{d}s + \int_{t-\epsilon}^u (b(V_{s_\delta}^\epsilon) - b(V_s^\epsilon))\mathrm{d}s.$$

Then for $u \in [t-\epsilon, t]$,

$$\begin{aligned}|Y_u - V_u^\epsilon| \leqslant& \int_{t-\epsilon}^u |b(Y_s) - b(V_s^\epsilon)|\mathrm{d}s + \int_{t-\epsilon}^u |b(V_{s_\delta}^\epsilon) - b(V_s^\epsilon)|\mathrm{d}s \\ &+ \int_{t-\epsilon}^t \int_{\mathbb{R}^d}\int_0^{k(Y_{s-},z)} g(Y_{s-},z)N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) \\ \leqslant& c_1\epsilon \sup_{s\in[t-\epsilon,u]} |Y_s - V_s^\epsilon|^{\theta_1} + c_1\epsilon \sup_{s\in[t-\epsilon,u]} |V_{s_\delta}^\epsilon - V_s^\epsilon|^{\theta_1} \\ &+ c_2\int_{t-\epsilon}^t \int_{\mathbb{R}^d}\int_0^{\lambda_2} |z|N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s).\end{aligned}$$

We can get

$$\mathbf{E}[R_{t,\epsilon}|^p] := \mathbf{E}\left[\left|\int_{t-\epsilon}^t \int_{\mathbb{R}^d}\int_0^{\lambda_2} |z|N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right|^p\right] \leqslant C\epsilon^{p/\alpha}$$

with the similar argument proving (6.59). Setting $S_{t,\epsilon} = \sup_{s\in[t-\epsilon,t]}|Y_s - V_s^\epsilon|$ and using that $b \in C^{\theta_1}(\mathbb{R}^d)$ and that $|V_s^\epsilon - V_{s_\delta}^\epsilon| \leqslant C\delta$, we see that

$$S_{t,\epsilon} \leqslant C(\epsilon S_{t,\epsilon}^{\theta_1} + \epsilon\delta^{\theta_1} + R_{t,\epsilon}) = C(\epsilon S_{t,\epsilon}^{\theta_1} + \epsilon^{\frac{1}{1-\theta_1}} + R_{t,\epsilon}).$$

Choosing ϵ sufficient small and using the Young inequality, we have $S_{t,\epsilon} \leqslant C\epsilon^{\frac{1}{1-\theta_1}} + \frac{\theta_1}{2}S_{t,\epsilon} + CR_{t,\epsilon}$. Thus,

$$S_{t,\epsilon} \leqslant CR_{t,\epsilon} + C\epsilon^{\frac{1}{1-\theta_1}}. \tag{6.68}$$

We finally recall that $Y_t^\epsilon = V_t^\epsilon + \int_{t-\epsilon}^t \int_{\mathbb{R}^d}\int_0^{k(Y_{t-\epsilon},z)} g(Y_{t-\epsilon},z)N(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s) = Y_{t-\epsilon} + \int_{t-\epsilon}^t b(V_s^\epsilon)\mathrm{d}s +$

$\int_{t-\epsilon}^t (b(V_{s_\delta}^\epsilon) - b(V_s^\epsilon))\mathrm{d}s + \int_{t-\epsilon}^t \int_{\mathbb{R}^d} \int_0^{k(Y_{t-\epsilon},z)} g(Y_{t-\epsilon}, z) N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s)$ so that

$$
\begin{aligned}
|Y_t - Y_t^\epsilon| \leqslant & \int_{t-\epsilon}^t |b(Y_s) - b(V_s^\epsilon)|\mathrm{d}s \\
& + \Big| \int_{t-\epsilon}^t \int_{\mathbb{R}^d} \int_0^\infty |\mathbf{1}_{[0,k(Y_{s-},z)]}(r) g(Y_{s-}, z) - \mathbf{1}_{[0,k(Y_{t-\epsilon},z)]}(r) g(Y_{t-\epsilon}, z)| N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) \Big| \\
& + \int_{t-\epsilon}^t |b(V_{s_\delta}^\epsilon) - b(V_s^\epsilon)|\mathrm{d}s =: I_{t,\epsilon} + J_{t,\epsilon} + K_{t,\epsilon}.
\end{aligned}
$$

First, by (6.68)

$$
I_{t,\epsilon} \leqslant C \int_{t-\epsilon}^t |Y_s - V_s^\epsilon|^{\theta_1} \mathrm{d}s \leqslant C(\epsilon R_{t,\epsilon}^{\theta_1} + \epsilon^{\frac{1}{1-\theta_1}}),
$$

thanks to the fact $\mathbf{E}|R_{t,\epsilon}|^p \leqslant C\epsilon^{p/\alpha}$,

$$
\mathbf{E}[|I_{t,\epsilon}|^p] \leqslant C[\epsilon^{\frac{p}{1-\theta_1}} + \epsilon^p \mathbf{E}(R_{t,\epsilon}^{p\theta_1})] \leqslant C[\epsilon^{\frac{p}{1-\theta_1}} + \epsilon^{p(1+\frac{\theta_1}{\alpha})}].
$$

Next for $J_{t,\epsilon}$, by the same way of dealing with (6.66), we have

$$
\begin{aligned}
\mathbf{E}[|J_{t,\epsilon}|^p] \leqslant & C\mathbf{E}\left[\left| \int_{t-\epsilon}^t \int_{|z|\leqslant \epsilon^{\frac{1}{\alpha}}} \int_0^{\lambda_2} \big|Y_{s-} - Y_{t-\epsilon}\big|^{2\theta_2} |z|^2 N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) \right|^{\frac{p}{2}} \right] \\
& + C\mathbf{E}\left(\int_{t-\epsilon}^t \int_{|z|>\epsilon^{\frac{1}{\alpha}}} \int_0^{\lambda_2} \big|Y_{s-} - Y_{t-\epsilon}\big|^{p\theta_2} |z|^p N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) \right) \\
& + C \left| \mathbf{E} \int_{t-\epsilon}^t \int_{|z|\leqslant \epsilon^{\frac{1}{\alpha}(1+\theta_3)}} \int_{k(Y_{t-\epsilon},z)}^{k(Y_{s-},z)} |z|^2 N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) \right|^{\frac{p}{2}} \\
& + C\mathbf{E}\left(\int_{t-\epsilon}^t \int_{|z|>\epsilon^{\frac{1}{\alpha}(1+\theta_3)}} \int_{k(Y_{t-\epsilon},z)}^{k(Y_{s-},z)} |z|^p N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) \right) \\
\leqslant & C\mathbf{E}\left[\sup_{s\in[t-\epsilon,t]} |Y_{s-} - Y_{t-\epsilon}|^{p\theta_2} \left(\int_{t-\epsilon}^t \int_{|z|\leqslant \epsilon^{\frac{1}{\alpha}}} \int_0^{\lambda_2} |z|^2 N(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s) \right)^{\frac{p}{2}} \right] \\
& + C\left(\int_{t-\epsilon}^t \int_{|z|>\epsilon^{\frac{1}{\alpha}}} \mathbf{E}\big|Y_{s-} - Y_{t-\epsilon}\big|^{p\theta_2} |z|^p \nu(\mathrm{d}z)\mathrm{d}s \right) \\
& + C\left| \int_{t-\epsilon}^t \int_{|z|\leqslant \epsilon^{\frac{1+\theta_3}{\alpha}}} (\mathbf{E}\,|Y_{s-} - Y_{t-\epsilon}|^{\theta_3} \wedge 1) |z|^2 \nu(\mathrm{d}z)\mathrm{d}s \right|^{\frac{p}{2}} \\
& + C\left(\int_{t-\epsilon}^t \int_{|z|>\epsilon^{\frac{1+\theta_3}{\alpha}}} \mathbf{E}(|Y_{s-} - Y_{t-\epsilon}|^{\theta_3} \wedge 1) |z|^p \nu(\mathrm{d}z)\mathrm{d}s \right) \\
\overset{(6.55),(6.56)}{\leqslant} & C\epsilon^{\frac{p}{\alpha}(1+\theta_2)} + C\epsilon^{\frac{p}{\alpha}(1+\theta_3)}.
\end{aligned}
$$

Finally, since $b \in C^{\theta_1}(\mathbb{R}^d)$ and since $|V_s^\epsilon - V_{s_\delta}^\epsilon| \leqslant C\delta$, we have $K_{t,\epsilon} \leqslant C\epsilon\delta^{\theta_1} = C\epsilon^{\frac{1}{1-\theta_1}}$ a.s., whence $\mathbf{E}[|K_{t,\epsilon}|^p] \leqslant C\epsilon^{\frac{p}{1-\theta_1}}$. Thus, we get (6.64) for $\alpha \in (0,1)$. The proof for $\alpha = 1$ is similar, so we omit it here. □

Now we are going to prove the regularity of the density of the process Y_t defined as in (6.54). We first give the following lemma about the regularity of Lévy processes.

Lemma 6.29. *Suppose Z_t is a Lévy process with Lévy measure ν, ν satisfies* **Assumption** *7. Let p_t^Z denote the density of Z_t, then for any $s \geqslant 0$, $q \in [1, \infty]$ and $t \in (0, 1)$,*

$$\|p_t^Z\|_{B^s_{q,\infty}} \leqslant Ct^{-(s+d/q')/\alpha}, \tag{6.69}$$

where $C = C(s, d, \alpha)$, $\frac{1}{q'} = 1 - \frac{1}{q}$.

Proof. Notice that

$$\|f\|_{B^s_{q,\infty}} = \sup_{j\geqslant -1} 2^{js}\|\Delta_j f\|_q \leqslant \left\|(\sum_{j\geqslant -1} |\Delta_j f|^2)^{1/2}\right\|_q \asymp \|f\|_{H^s_q},$$

where H^s_q is the Bessel potential space. By interpolation theorem, we only need to prove

$$\sup_{|\alpha|=k} \|\partial^\alpha p_t\|_q \leqslant C(k, d, \alpha)t^{-(k+d/q')/\alpha}, \quad k \in \mathbb{N},$$

and the above inequality is a simple consequence of [56, Proposition 2.3] and [15, Lemma 1.3 and Lemma 3.3]. So we complete our proof. □

Lemma 6.30. *Suppose a, g satisfy (6.63), $\theta_1 > 1-\alpha$ if $\alpha \in (0, 1)$ and $|g(y, z)| \geqslant c_2'|z|$ for some $c_2' > 0$, k satisfies* **Assumption** *5 with Λ_i and ϑ replaced by λ_i and θ_3, respectively. Then Y_t has a density p_t^Y and $p_t^Y \in B^\gamma_{q,\infty}$ with γ, q satisfying*

$$0 < \gamma < (1 \wedge \alpha)(\alpha\theta_0 - 1), \quad 1 \leqslant q < \frac{d}{d + \gamma - (1 \wedge \alpha)(\alpha\theta_0 - 1)}, \tag{6.70}$$

where θ_0 is the same number in Lemma 6.28.

Proof. Recalling that $\mathcal{C}_R = R \cdot \mathcal{C}$, for $\gamma > 0$ and $q \in [1, \infty]$ define

$$\mathcal{S}^{-\gamma}_{q,j} := \left\{\varphi \in \mathscr{S}(\mathbb{R}^d) : \hat{\varphi} \in \mathcal{C}_{2^j}, \|\varphi\|_q \leqslant 2^{\gamma j}\right\}.$$

Choose $\varphi \in \mathcal{S}^{-\gamma}_{q',j}$, take the constructed process $V_t^\epsilon, Y_t^\epsilon$ from Lemma 6.28,

$$Y_t^\epsilon = V_t^\epsilon + \int_{t-\epsilon}^t \int_{\mathbb{R}^d} \int_0^\infty \mathbf{1}_{[0,k(Y_{t-\epsilon},z)]}(r)g(Y_{t-\epsilon}, z)N^{(\alpha)}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s).$$

By trangale inequality,

$$|\mathbf{E}\varphi(Y_t)| \leqslant |\mathbf{E}\varphi(Y_t^\epsilon)| + |\mathbf{E}\varphi(Y_t) - \mathbf{E}\varphi(Y_t^\epsilon)| =: I_1^\epsilon(\varphi) + I_2^\epsilon(\varphi).$$

Define

$$Z_t^y := \int_0^t\int_{\mathbb{R}^d} \int_0^\infty \mathbf{1}_{[0,k(y,z)]}(r)g(y, z)N^{(\alpha)}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s), \quad y \in \mathbb{R}^d.$$

Then Z^y_t is a Lévy process with Lévy measure $\nu_y = \nu \circ [k(y,\cdot)g(y,\cdot)]^{-1}$. Under our assumptions, one can easily check that ν_y satisfies **Assumpiton** 7. For $I^\epsilon_1(\varphi)$, recall that $V^\epsilon_t \in \mathscr{F}_{t-\epsilon}$, we get

$$
\begin{aligned}
I^\epsilon_1(\varphi) =& |\mathbf{E}\,[(\mathbf{E}\,\varphi(Y^\epsilon_t)|\mathscr{F}_{t-\epsilon})]| \\
=& \left|\mathbf{E}\left[\mathbf{E}\left(\varphi\left(V^\epsilon_t + \int_{t-\epsilon}^t\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(Y_{t-\epsilon},z)]}(r)g(Y_{t-\epsilon},z)N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right)\Big|\mathscr{F}_{t-\epsilon}\right)\right]\right| \\
=& \left|\mathbf{E}\left[\mathbf{E}\left(\varphi\left(u + \int_{t-\epsilon}^t\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(y,z)]}(r)g(y,z)N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right)\Big|_{u=V^\epsilon_t, y=Y_{t-\epsilon}}\right)\right]\right|.
\end{aligned}
$$

Define $\tau_u\varphi(\cdot) := \varphi(\cdot + u)$ for $u \in \mathbb{R}^d$. By Lemma 6.29 and Bernstein's inequality, for $q' = \frac{q}{q-1}$ and $s > \gamma$

$$
\begin{aligned}
I^\epsilon_1(\varphi) &\leqslant \sup_{u\in\mathbb{R}^d} \mathbf{E}\,\varphi\left(u + \int_{t-\epsilon}^t\int_{\mathbb{R}^d}\int_0^\infty \mathbf{1}_{[0,k(y,z)]}(r)g(y,z)N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z,\mathrm{d}s)\right) \\
&= \sup_{u\in\mathbb{R}^d} \mathbf{E}\tau_u\varphi(Z^y_\epsilon) \leqslant C\|\varphi\|_{B^{-s}_{q',1}}\|p^{Z^y}_\epsilon\|_{B^s_{q,\infty}} \\
&\leqslant C2^{(\gamma-s)j}\epsilon^{-\frac{1}{\alpha}(s+\frac{d}{q'})}.
\end{aligned} \tag{6.71}
$$

Choose $p \in (0, 1\wedge\alpha)$, by Bernstein's inequality and Lemma 6.28,

$$
\begin{aligned}
I^\epsilon_2(\varphi) \leqslant& \|\varphi\|_{C^p}\mathbf{E}|Y^\epsilon_t - Y_t|^p \leqslant C2^{(p+\frac{d}{q'})j}\|\varphi\|_{q'}\,\epsilon^{\theta_0 p} \\
\leqslant& C2^{(p+\gamma+\frac{d}{q'})j}\epsilon^{\theta_0 p}.
\end{aligned} \tag{6.72}
$$

where θ_0 keeps the same as in (6.64). Notice that under our assumptions, $\alpha\theta_0 > 1$, for any

$$p \in (0, 1\wedge\alpha), \quad 0 < \gamma < (\alpha\theta_0 - 1)p,$$

we can choose s, q, ϵ such that

$$q < \frac{d}{d+\gamma-(\alpha\theta_0-1)p}, \quad s = \frac{\alpha\theta_0 p\gamma + d(p+\gamma+d/q')/q'}{\alpha\theta_0 p - p - \gamma - d/q'}, \quad \epsilon = 2^{\frac{\alpha(\gamma-s)j}{s+d/q'}}. \tag{6.73}$$

Then combine (6.71), (6.72) and (6.73), we get

$$|\mathbf{E}\varphi(Y_t)| \leqslant I^\epsilon_1(\varphi) + I^\epsilon_2(\varphi) \leqslant C, \tag{6.74}$$

where C only depends on $d, \alpha, \theta_i, \lambda_i, c_i, \gamma, p, q$. When $\alpha \in [1,2)$, notice that p can infinitely approach 1, so we have

$$0 < \gamma < (\alpha\theta_0 - 1), \quad 1 \leqslant q < \frac{d}{d+1+\gamma-\alpha\theta_0}.$$

When $\alpha \in (0,1)$, p can infinitely approach α, so

$$0 < \gamma < \alpha(\alpha\theta_0 - 1), \quad 1 \leqslant q < \frac{d}{d+\alpha+\gamma-\alpha^2\theta_0}.$$

For any $\varphi \in B^{-\gamma}_{q',1}$ and $j \geqslant -1$, define $\varphi_j = \frac{\Delta_j \varphi}{2^{\gamma j}\|\Delta_j \varphi\|_{q'}}$. Notice that $\varphi_j \in \mathcal{S}^{-\gamma}_{q',j}$, by (6.74), we obtain

$$|\mathbf{E}\varphi(Y_t)| \leqslant \sum_{j\geqslant -1} |\mathbf{E}\Delta_j\varphi(Y_t)| \leqslant \sum_{j\geqslant -1} |\mathbf{E}\varphi_j(Y_t)| \cdot 2^{\gamma j}\|\Delta_j\varphi\|_{q'} \leqslant C\|\varphi\|_{B^{-\gamma}_{q',1}}.$$

By duality, $p_t^Y \in B^{\gamma}_{q,\infty}$. □

Now suppose $\kappa(x,z)$ satisfies ($\mathbf{H}_1$)-($\mathbf{H}_3$) and $\max\{0,(1-\alpha)\} < \vartheta < 1$, $\beta \in (1-\alpha, \vartheta)$ when $\alpha \in (0,1]$, $\beta \in (-(\frac{\alpha-1}{2} \wedge \vartheta), 0]$ when $\alpha \in (1,2)$, and $b \in \mathscr{C}^\beta$. By Theorem 6.1, we can fix λ sufficient large such that $u \in \mathscr{C}^{\alpha+\beta}$ is the unique solution to the following resolvent equation in the distribution sense

$$\lambda u - \mathscr{L}^\alpha_{\kappa,b} u = b,$$

and

$$\|\nabla u\|_{L^\infty(\mathbb{R}^d)} \leqslant \frac{1}{2}.$$

Define $\Phi(x) =: u(x) + x$, then Φ is a diffeomorphism.

Proposition 6.31. *Under the same conditions as in Corollary 6.2, the process $Y_t := \Phi(X_t)$ satisfies the following SDE*

$$Y_t = Y_0 + \int_0^t a(Y_s)\mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} \int_0^\infty g(Y_{s-}, z)\mathbf{1}_{[0,k(Y_{s-},z)]}(r) N^{(\alpha)}(\mathrm{d}r, \mathrm{d}z, \mathrm{d}s),$$

where X_t is the weak solution to (6.53),

$$a(y) := \lambda u(\Phi^{-1}(y)), \quad k(y,z) := \kappa(\Phi^{-1}(y), z) \tag{6.75}$$

and

$$g(y,z) = \Phi(\Phi^{-1}(y)+z) - y = u(\Phi^{-1}(y)+z) + z - u(\Phi^{-1}(y)). \tag{6.76}$$

Furthermore, we have $a \in \mathscr{C}^{\alpha+\beta}$,

$$c_2^{-1}|z| \leqslant |g(y,z)| \leqslant c_2|z|, \quad |g(y_1,z) - g(y_2,z)| \leqslant c_3|y_1-y_2|^{\alpha+\beta-1}|z| \tag{6.77}$$

and k satisfies ($\mathbf{H}_1$)-($\mathbf{H}_3$) with the same ϑ as κ.

Proof. With the similar argument showed in [5, Proposition 2.7], applying Itô's formula to $\Phi(x) = u(x) + x$ with respect to the process X_t, we get the desired conclusion. □

Now we are in the position of proving Theorem 6.4.

Proof of Theorem 6.4. For $\alpha \in (0,1]$, letting $a = b$, $g(y,z) = z$, $k = \kappa$, we have $\theta_1 = \beta$, θ_2 can infinitely approach 1, $\theta_3 = \vartheta$. By Lemma 6.30, we have $\alpha\theta_0 = \alpha + \beta$ and $p_t^X \in B^{\gamma}_{q,\infty}$ with

$$0 < \gamma < \alpha(\alpha + \beta - 1), \quad 1 \leqslant q < \frac{d}{d + \gamma - \alpha(\alpha + \beta - 1)}.$$

For $\alpha \in (1,2)$, by Proposition 6.31, $Y_t = \Phi(X_t)$ satisfies (6.54) and in this case the index θ_1 can be taken infinitely approach 1, $\theta_1 = \alpha + \beta - 1$ and $\theta_3 = \vartheta$. Therefore, by Lemma 6.30, $p_t^Y \in B^{\gamma}_{q,\infty}$ with

$$0 < \gamma < (\alpha + \beta - 1) \wedge \tfrac{\vartheta}{\alpha}, \quad 1 \leqslant q < \frac{d}{d + \gamma - (\alpha + \beta - 1) \wedge \frac{\vartheta}{\alpha}}.$$

This implies that there also exists a density p_t^X of the distribution of X_t such that $p_t^X = p_t^Y \circ \Phi \cdot \det(\nabla\Phi)$ and $p_t^X \in B^{\gamma}_{q,\infty}$. Since the martingale solution $\mathbb{P}$ corresponding to SDE (6.53) can be denoted by $\mathbb{P} = \mathbf{P} \circ X$, we get the desired result. □

Last we point out that Corollary 6.5 is a consequence of Corollary 6.2, Theorem 6.4 and Proposition 4.9.19 of [20].

A Appendix

A.1 Khasminskii's lemma

Lemma A.1. *([51, P. 1 Lemma 1.1.]) Let $\{\beta(t)\}_{t\in[0,T]}$ be a nonnegative measurable $\{\mathcal{F}_t\}-$adapted process. Assume that for all $0 \leqslant s \leqslant t \leqslant T$,*

$$E\left(\int_s^t \beta(r)dr\bigg|_{\mathcal{F}_s}\right) \leqslant \Gamma(s,t),$$

where $\Gamma(s,t)$ is a nonrandom interval function satisfying the following conditions:
(i)$\Gamma(t_1,t_2) \leqslant \Gamma(t_s,t_4)$ if $(t_1,t_2) \subset (t_3,t_4)$;
(ii)$\lim_{h\downarrow 0}\sup_{0\leqslant s<t\leqslant T,|t-s|\leqslant h}\Gamma(s,t) = \lambda$, $\lambda \geqslant 0$. Then for any arbitrary real $\kappa < \lambda^{-1}$ (if $\lambda = 0$, then $\lambda^{-1} = \infty$),

$$Eexp\left\{\kappa\int_0^T \beta(r)dr\right\} \leqslant C = C(\kappa,\Gamma,T) < \infty.$$

A.2 Non-explosion lemma

Lemma A.2. *X_t is a processes in $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},P)$, $\tau =: \inf\{t \geqslant 0 : |X_t| = \infty\}$. Process X_t is non-explosive (i.e. $\tau = \infty$ a.s.) if for any $t > 0$ one of the following conditions holds:*
(i) $E|X_{t\wedge\tau}| \leqslant C(t)$.
(ii) $E|X_t| \leqslant C(t)$.

Proof. $\Omega =: \{\omega : \tau(\omega) = \infty\}$, $\Omega_n =: \{\omega : \tau(\omega) > n\}$. Then $\Omega = \cap_{n=1}^{\infty}\Omega_n$, and $\Omega_n = \big\{\omega : |X_{n\wedge\tau(\omega)}(\omega)| < \infty\big\} = \{\omega : |X_n(\omega)| < \infty\}$. Since T is arbitrary in $(0,\infty)$, for any $N \in \mathbb{N}$, $n \in [0,N]$, from condition (i) we get $E|X_{n\wedge\tau}| \leqslant C(N)$, it implies $|X_{n\wedge\tau}| < \infty$ *a.s.*, i.e. $P(\Omega_n) = 1$. Hence $P(\Omega) = 1$, which implies $\tau = \infty$ *a.s.*. (i) is proved.

For the second one, from condition (ii) we get $E|X_n| \leqslant C(N)$, so $|X_n| < \infty$ *a.s.*, then $P(\Omega_n) = 1$, $P(\Omega) = 1$, $\tau = \infty$ *a.s.*. □

A.3 Girsanov transformation

Lemma A.3. *Let σ and $b^{(i)}$, $i = 1,2$ satisfy the requirements in the beginning of subsection 3.2 and let $|b^{(1)} - b^{(2)}| \leqslant \overline{b}$, where $\overline{b} \in \mathbb{L}_p^q$ with p, q satisfying (5.2). Let $(X_t^{(i)}, W_t^{(i)})$ satisfy:*

$$X_t^{(i)} = \int_0^t b^{(i)}(s,X_s^{(i)})ds + \int_0^t \sigma(s,X_s^{(i)})dW_s^{(i)}.$$

Then for any bounded Borel functions $f(x)$, given on $\mathcal{C} =: \mathcal{C}([0,\infty),\mathbb{R}^d)$ we have

$$Ef(X_\cdot^{(2)}) = Ef(X_\cdot^{(1)})\overline{\rho}_\infty$$

if

$$E\exp\Big(\frac{1}{2}\int_0^\infty (\Delta b^*(\sigma\sigma^*)^{-1}\Delta b)(s,X_s^{(1)})ds\Big) < \infty, \tag{7.1}$$

where $\Delta b(t,X_t^{(1)}) =: b^{(2)}(t,X_t^{(1)}) - b^{(1)}(t,X_t^{(1)})$ *and*

$$\overline{\rho}_t =: \exp(\int_0^t \Delta b^*(\sigma^*)^{-1}(s,X_s^{(1)})dW_s^{(1)} - \frac{1}{2}\int_0^t (\Delta b^*(\sigma\sigma^*)^{-1}\Delta b)(s,X_s^{(1)})ds).$$

Proof. Theorem 6.1 in [47] says if (7.1) (Novikov condition) holds, then $(\overline{\rho}_t)_{t\geqslant 0}$ is a $(\mathcal{F}_t)_{t\geqslant 0}$ martingale. Let $\hat{P} = \overline{\rho}_\infty P$, then $\hat{P}$ is also a probability on $(\Omega,\mathcal{F})$. By Theorem 4.1 in [33],

$$\hat{W}(t) = W^{(1)}(t) - \int_0^t \sigma^{-1}(s,X_s^{(1)})\Delta b(s,X_s^{(1)})ds$$

is a $\mathcal{F}_t$ Brownian motion on the probability space $(\Omega,\mathcal{F},\hat{P})$. So we can wirte

$$\begin{aligned}
X_t^{(1)} &= \int_0^t b^{(1)}(s,X_s^{(1)})ds + \int_0^t \sigma(s,X_s^{(1)})d\hat{W}_s + \int_0^t \sigma(s,X_s^{(1)})\sigma^{-1}(s,X_s^{(1)})\Delta b(s,X_s^{(1)})ds\\
&= \int_0^t b^{(1)}(s,X_s^{(1)})ds + \int_0^t \sigma(s,X_s^{(1)})d\hat{W}_s + \int_0^t \Delta b(s,X_s^{(1)})ds\\
&= \int_0^t b^{(2)}(s,X_s^{(1)})ds + \int_0^t \sigma(s,X_s^{(1)})d\hat{W}_s.
\end{aligned}$$

This implies that $(X_t^{(1)},\hat{W}(t))$ is a solution on the probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\hat{P})$ to the SDE

$$X_t^{(2)} = \int_0^t b^{(2)}(s,X_s^{(2)})ds + \int_0^t \sigma(s,X_s^{(2)})dW_s^{(2)}. \tag{7.2}$$

From Lemma 5.5 we know that the solution to SDE (7.2) is unique, hence for any bounded Borel functions $f(x)$, given on $\mathcal{C} =: \mathcal{C}([0,\infty),\mathbb{R}^d)$ we have

$$Ef(X_\cdot^{(2)}) = \hat{E}f(X_\cdot^{(1)}) = E\overline{\rho}_\infty f(X_\cdot^{(1)}).$$

□

A.4 Urysohn Lemma

For the convenience of reading, we put the $\mathcal{C}^\infty$ Urysohn Lemma here.

Lemma A.4. *([24, 8.18]) If $K \subset \mathbb{R}^n$ is compact and U is an open set containing K, there exists smooth function f such that $0 \leqslant f \leqslant 1$, $f = 1$ on K, and supp$(f) \subset U$.*

A.5 Equivalence between martingale solutions and weak solutions

Consider the following SDE:

$$dX_t = \int_{\mathbb{R}^d}\int_0^\infty z\mathbf{1}_{[0,\kappa(X_{t-},z))}(r)N^{(\alpha)}(\mathrm{d}r,\mathrm{d}z;\mathrm{d}t) + b(X_t)dt, \quad X_0 = x \in \mathbb{R}^d. \tag{7.3}$$

Where $N^{(\alpha)}(dr,dz;dt)$ is defined in (6.36). $b \in \mathscr{C}^\beta(\mathbb{R}^d)$ with $\beta \in \mathbb{R}$.

Recall the definition in 6.25, we could not get that the solution X is a semimartingale. Since condition 3 shows the quadratic variation of A is 0, but A may not be of finite variation. In order to handle with this case, we introduce a more general class of processes called Dirichlet processes. We begin with the definition of the processes of zero energy from [25]

Definition A.5. *We say that a continuous adapted process $(A_t)_{t\in[0,T]}$ is a process of zero energy if $A_0 = 0$ and*

$$\lim_{\delta\to 0}\sup_{|\pi_T|<\delta} E\Big(\sum_{t_i\in\pi_T}|A_{t_{i+1}} - A_{t_i}|^2\Big) = 0$$

where π_T denotes a finite partition of $[0,T]$ and $|\pi_T|$ denotes the mesh size of the partition.

Definition A.6 ([25])**.** *We say that an adapted process $(X_t)_{t\in[0,T]}$ is a Dirichlet process if*

$$X_t = M_t + A_t \tag{7.4}$$

where M is a square-integrable martingale and A is an adapted process of zero energy.

Let $p \geqslant 1$ and $\beta > 0$. For a stochastic process $(X_t)_{t\in[0,T]}$ and $T > 0$, we define

$$\mathcal{H}_T^{r,p}(X) := \|X_0\|_{L^p(\Omega)} + \sup_{s\neq t,s,t\in[0,T]}\frac{\|X_t - X_s\|_{L^p(\Omega)}}{|t-s|^r}.$$

Lemma A.7 ([78], Lemma 3.12, [5], Lemma 3.10)**.** *Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a bounded continuous function with a bounded continuous derivative. Let X be a Dirichlet process with decomposition (7.4). Let p, q, $r \in [1,\infty)$ with $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Suppose that for any $T > 0$, there are γ, $\eta \in (0,1]$ with $\gamma + \eta > 1$ such that*

$$\mathcal{H}_T^{\eta,p}(f(X)) < \infty, \quad \mathcal{H}_T^{\gamma,q}(A) < \infty.$$

For $n \in \mathbb{Z}_+$, for $t \in [0,T]$, $t_n^k := k2^{-n}t$, the sequence of partial sums $I_n := \sum_{i=0}^{2^n-1} f(X_{t_n^i})$ $(A_{t_n^{i+1}} - A_{t_n^i})$ converges in $L^r(\Omega)$ and the limit is $\int_0^t f(X_s)dA_s$. Moreover, there is a constant $C > 0$ depending only on η, γ and T such that for all $t \in [0,T]$,

$$\|I_n - \int_0^t f(X_s)dA_s\|_{L^r(\Omega)} \leqslant C\mathcal{H}_T^{\eta,p}(f(X))\mathcal{H}_T^{\gamma,q}(A)2^{-n(\gamma+\eta-1)}, \tag{7.5}$$

and for all $0 \leqslant t' < t \leqslant T$,

$$\| \int_{t'}^{t} f(X_s) dA_s \|_{L^r(\Omega)} \leqslant C \mathcal{H}_T^{\eta,p}(f(X)) \mathcal{H}_T^{\gamma,q}(A)(t-t')^{\gamma}. \tag{7.6}$$

Further, let $(f_n)_{n\in\mathbb{N}}$ *be a sequence of functions on* $\mathbb{R}^d$ *that a uniformly bounded, continuous, and have a bounded continuous derivative such that* $\mathcal{H}_T^{\eta,p}(f_n(X)) < \infty$ *for all* $n \in \mathbb{N}$*. Let* $(b_n)_{n\in\mathbb{N}}$ *be a sequence of bounded continuous functions. Define* $A_t^{b_n} := \int_0^t b_n(X_s)ds$ *for* $t \in [0,T]$*. If* $\mathcal{H}_T^{\gamma,q}(A^{b_n}) < \infty$*, and if* $A_t^{b_n}$ *converges in probability to* A_t *for each* $t \in [0,T]$*, then for any* $t \in [0,T]$ *we have*

$$\int_0^t f_n(X_s) dA_s - \int_0^t f_n(X_s) b_n(X_s) ds \to 0, \ \textit{in probability as } n \to \infty. \tag{7.7}$$

Now we are going to show the equivalence.

Theorem A.8. *Suppose*

1. $\alpha \in (0,1]$, $\kappa(x,z)$ *satisfies* (**H**$_1$)-(**H**$_3$) *with* $\max\{0,(1-\alpha)\} < \vartheta < 1$*, and* $b \in \mathscr{C}^{\beta}$ *with* $\beta \in (0,\vartheta)$
2. $\alpha \in (1,2)$, $\kappa(x,z)$ *satisfies* (**H**$_1$)-(**H**$_3$) *and* $b \in \mathscr{C}^{\beta}$ *with* $\beta \in (-(\frac{\alpha-1}{2} \wedge \vartheta), 0]$.

Let $\mathbb{P} \in \mathcal{P}(\mathbb{D})$*, then* $\mathbb{P} \in \mathcal{M}_{k,b}^{\beta}(x)$ *if and only if there is a weak solution* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P}, X, N, A)$ *so that* $\mathbf{P} \circ X^{-1} = \mathbb{P} \in \mathcal{K}_{\beta}(\mathbb{D})$.

Proof. We only show the case when $\alpha \in (1,2)$. With the similar argument we get the result for $\alpha \in (0,1]$.

If $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P}, X, N, A)$ is a weak solution of SDE (7.3) satisfying $\mathbf{P} \circ X^{-1} = \mathbb{P} \in \mathcal{K}_{\beta}(\mathbb{D})$, then from Definition 7.4 we know that X is a Dirichlet process. For any $f \in \mathcal{C}^{\infty}$, by applying Itô's formula for the Dirichlet process X we get

$$\begin{aligned} f(X_t) = f(x) &+ \int_0^t \mathcal{L}_k^{\alpha} f(X_s) ds + \int_0^t \nabla f(X_s) dA_s \\ &+ \int_0^t \int_{\mathbb{R}^d} \int_0^{\infty} \Big(f(X_{s-} + 1_{[0,k(X_{s-},z)](r)} z) - f(X_s) \Big) \tilde{N}(dr,dz;ds) \end{aligned}$$

where the term $\int_0^t \nabla f(X_s) dA_s$ is in the sense of Lemma A.7. In order to show $\mathbf{P} \circ X^{-1} \in \mathcal{M}_{k,b}^{\beta}(x)$, by Proposition 4.2 we only need to prove that for any $t > 0$

$$\int_0^t \nabla f(X_s) dA_s = A_t^{b \cdot \nabla f}, \quad \mathbf{P} - a.s., \tag{7.8}$$

where $A_t^{b_n} := \int_0^t b_n(X_s) ds$ and

$$A_t^{b \cdot \nabla f} = \lim_{n\to\infty} \int_0^t \nabla f(X_s) dA_s^{b_n} = \lim_{n\to\infty} \int_0^t (b_n \cdot \nabla f)(X_s) ds \text{ in probability.}$$

Here $b_n := b * \rho_n$. Since we have $A^{b_n} \to A^b$ in the sense of *u.c.p.*, by (7.7) we have

$$\int_0^t b_n \cdot \nabla f(X_s) ds \to \int_0^t \nabla f(X_s) dA_s \text{ in probability as } n \to \infty,$$

which yields (7.8).

Now we are going to prove the another way. Suppose that $\mathbb{P} \in \mathcal{M}^{\beta}_{k,b}(x)$ satisfies (6.44), if we take $X_t(\omega) = \omega(t)$, $A_t = A^b_t$, then

$$\mathbb{E}|A^b_t - A^b_s|^2 \leqslant C|t-s|^{1+\gamma},$$

which shows that the condition (ii) in Definition 6.25 is fulfilled. Then by [39, Theorem 2.3], we get the desired result.

□

A.6 The Sobolev embedding theorem in mixed-norm spaces

In order to show Sobolev embedding theorem in Mixed-norm space, we first prove the following lemma.

Lemma A.9. *Let $\mathcal{C}$ be an annulus and B a ball. A constant C exists such that for any nonnegative integer k, any couple $(\mathbf{p}, \mathbf{q}) \in [1,\infty]^{2d}$ with $q_i \geqslant p_i \geqslant 1$, $1 \leqslant i \leqslant d$, and any function u of $L^{\mathbf{p}}$, we have*

$$Supp(\hat{u}) \subset \lambda B \Rightarrow \|D^k u\|_{L^{\mathbf{q}}} \stackrel{def}{=} \sup_{|\alpha|=k} \|\partial^\alpha u\|_{L^{\mathbf{q}}} \leqslant C^{k+1} \lambda^{k+(\frac{1}{p_1}+\cdots+\frac{1}{p_d})-(\frac{1}{q_1}+\cdots+\frac{1}{q_d})} \|u\|_{L^{\mathbf{p}}},$$

$$Supp(\hat{u}) \subset \lambda \mathcal{C} \Rightarrow C^{-k-1}\lambda^k \|u\|_{L^{\mathbf{p}}} \leqslant \|D^k u\|_{L^{\mathbf{p}}} \leqslant C^{k+1}\lambda^k \|u\|_{L^{\mathbf{p}}}.$$

Proof. Using a dilation of size λ, we can assume throughout the proof that $\lambda = 1$. Let ϕ be a smooth function defined on $\mathbb{R}^d$ with value 1 near B. As $\hat{u}(\xi) = \phi(\xi)\hat{u}(\xi)$, we have

$$\partial^\alpha u = \partial^\alpha g * u \text{ with } g = \mathcal{F}^{-1}\phi.$$

Applying Young's inequality we get

$$\|\partial^\alpha g * u\|_{L^{\mathbf{q}}} \leqslant \|\partial^\alpha g\|_{L^{\mathbf{r}}} \|u\|_{L^{\mathbf{p}}} \text{ with } \mathbf{r} = (r_1, \cdots, r_d), \quad \frac{1}{r_i} \stackrel{def}{=} -\frac{1}{p_i} + \frac{1}{q_i} + 1, \quad 1 \leqslant i \leqslant d.$$

And the first assertion follow via

$$\begin{aligned} \|\partial^\alpha g\|_{L^{\mathbf{r}}} &\leqslant \|\partial^\alpha g\|_{L^\infty} + \|\partial^\alpha g\|_{L^1} \\ &\leqslant C\|(1+|\cdot|^2)^d \partial^\alpha g\|_{L^\infty} \\ &\leqslant C\|(\mathbb{I}_{d\times d} - \Delta)^d((\cdot)^\alpha \phi)\|_{L^1} \\ &\leqslant C^{k+1}. \end{aligned}$$

To prove the second assertion, consider a function $\tilde{\phi} \in \mathcal{S}(\mathbb{R}^d \backslash \{0\})$ with value 1 on a neighborhood of $\mathcal{C}$. By the fact that there exists a family of integers $(A_\alpha)_\alpha \in \mathbb{N}^d$ such that for $\xi \in \mathbb{R}^d$,

$$|\xi|^{2k} = \sum_{1 \leqslant j_1, \cdots, j_k \leqslant d} \xi_{j_1}^2 \cdots \xi_{j_k}^2 = \sum_{|\alpha|=k} A_\alpha (i\xi)^\alpha (-i\xi)^\alpha,$$

we get that

$$\hat{u}(\xi) = \sum_{|\alpha|=k} (i\xi)^\alpha v_\alpha(\xi) \quad \text{with } v_\alpha(\xi) \stackrel{def}{=} A_\alpha \frac{(-i\xi)^\alpha}{|\xi|^{2k}} \hat{u}(\xi).$$

Since $\hat{u} = \tilde{\phi}\hat{u}$, we deduce

$$u = \sum_{|\alpha|=k} g_\alpha * \partial^\alpha u \text{ with } g_\alpha \stackrel{def}{=} A_\alpha \mathcal{F}^{-1}(-i\xi)^\alpha |\xi|^{-2k} \tilde{\phi}(\xi),$$

and the result follows. □

Lemma A.10. *For* $\mathbf{p}, \mathbf{q} \in [1, \infty]^d$ *with* $p_1 \leqslant q_1, \cdots, p_d \leqslant q_d$. *Then for any real numbers* $s > (\frac{1}{p_1} + \cdots + \frac{1}{p_d}) - (\frac{1}{q_1} + \cdots + \frac{1}{q_d})$ *and* $r \in [1, \infty]$,

$$\|f\|_{B_{\mathbf{q},r}^{s-[(\frac{1}{p_1}+\cdots+\frac{1}{p_d})-(\frac{1}{q_1}+\cdots+\frac{1}{q_d})]}} \leqslant C\|f\|_{B_{\mathbf{p},r}^s}.$$

Proof. According to Littlewood-Paley theorem which is introduced in Section 6.2, it suffices to prove

$$\|\Delta_0 f\|_{L^{\mathbf{q}}} \leqslant C\|\Delta_0 f\|_{L^{\mathbf{p}}} \tag{7.9}$$

and

$$\|\Delta_j f\|_{L^{\mathbf{q}}} \leqslant C 2^{j[(\frac{1}{p_1}+\cdots+\frac{1}{p_d})-(\frac{1}{q_1}+\cdots+\frac{1}{q_d})]} \|\Delta_j f\|_{L^{\mathbf{p}}} \tag{7.10}$$

for all $j \in \mathbb{N}$, which can be obtained by applying the above Lemma A.9.

□

Lemma A.11. *Let* $\alpha > 0$ *and* $\mathbf{p} \in (1, \infty)^d$. *Then*

(1) $H_{\mathbf{p}}^\alpha(\mathbb{R}^d) \hookrightarrow L^{\mathbf{q}}$ *for all* $\mathbf{q} \in [\mathbf{p}, \mathbf{p}^*]$, $(\frac{1}{p_1^*} + \cdots + \frac{1}{p_d^*}) = (\frac{1}{p_1} + \cdots + \frac{1}{p_d}) - \alpha$, *when* $(\frac{1}{p_1} + \cdots + \frac{1}{p_d}) > \alpha$.

(2) $H_{\mathbf{p}}^\alpha(\mathbb{R}^d) \hookrightarrow \mathcal{C}(\mathbb{R}^d)$, *when* $(\frac{1}{p_1} + \cdots + \frac{1}{p_d}) < \alpha$. *Here* $\mathcal{C}(\mathbb{R}^d)$ *is equipped with the supremum norm.*

Proof. From [1, (1.2.29)], the Bessel potential space $H_{\mathbf{p}}^\alpha(\mathbb{R}^d)$ can also be defined by

$$H_{\mathbf{p}}^\alpha(\mathbb{R}^d) = \{f : f = G_\alpha * g, g \in L^{\mathbf{p}}(\mathbb{R}^d)\}, \quad \alpha \in \mathbb{R},$$

and

$$\|f\|_{H^\alpha_{\mathbf{p}}} = \|g\|_{L^{\mathbf{p}}},$$

where $G_\alpha(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \frac{e^{i(x,\xi)}}{(1+|\xi|^2)^{\alpha/2}}d\xi$ and $g := \mathcal{G}_{-\alpha}f = (I-\Delta)^{\alpha/2}f := G_{-\alpha} * f$. For $\alpha > 0$, by [1, (1.2.11)],

$$G_\alpha(x) = \frac{1}{(4\pi)^{\alpha/2}\Gamma(\alpha/2)}\int_0^\infty t^{\frac{\alpha-d}{2}} e^{-\frac{\pi|x|^2}{t}-\frac{t}{4\pi}}\frac{dt}{t}.$$

Then by Hölder's inequality,

$$|G_\alpha * g| \leqslant \|G_\alpha\|_{\mathbf{p}'}\|g\|_{\mathbf{p}}$$

if $(\frac{1}{p_1}+\cdots+\frac{1}{p_d}) < \alpha$ since in this case $G_\alpha \in L^{\mathbf{p}'}$ where $\frac{1}{p_1}+\frac{1}{p_1'} = 1, \cdots, \frac{1}{p_d}+\frac{1}{p_d'} = 1$. To be more precisely,

$$\begin{aligned}
\|G_\alpha\|_{\mathbf{p}'} =& \Big(\int_{\mathbb{R}}\cdots\Big(\int_{\mathbb{R}}|G_\alpha(x)|^{p_1'}dx_1\Big)^{\frac{p_2'}{p_1'}}\cdots dx_d\Big)^{\frac{1}{p_d'}}\\
=&\frac{1}{(4\pi)^{\alpha/2}\Gamma(\alpha/2)}\Big(\int_{\mathbb{R}}\cdots\Big(\int_{\mathbb{R}}|\int_0^\infty t^{\frac{\alpha-d}{2}}e^{-\frac{\pi|x|^2}{t}-\frac{t}{4\pi}}\frac{dt}{t}|^{p_1'}dx_1\Big)^{\frac{p_2'}{p_1'}}\cdots dx_d\Big)^{\frac{1}{p_d'}}\\
\leqslant&\frac{1}{(4\pi)^{\alpha/2}\Gamma(\alpha/2)}\int_0^\infty\Big(e^{-\frac{t}{4\pi}}t^{\frac{1}{2}(\frac{1}{p_1'}+\cdots+\frac{1}{p_d'}-d+\alpha)}\Big)\frac{dt}{t}\\
=&1
\end{aligned}$$

if $\frac{1}{p_1'}+\cdots+\frac{1}{p_d'}-d+\alpha > 0$, i.e. $(\frac{1}{p_1}+\cdots+\frac{1}{p_d}) < \alpha$. The continuity of $G_\alpha * g$ follows from the continuity of translation in the $L^{\mathbf{p}}$ norm. Then (2) is proved. □

References

[1] Adams D. and Hedberg L.: Function Spaces and Potential Theory. *Springer-Verlag Berlin Heidelberg New York*, (1991).

[2] Applebaum D.: Lévy Processes and Stochastic Calculus. *Cambridge University Press,* (2009).

[3] Antonić N. and Ivec I.: On the Hörmander-Mihlin theorem for mixed-norm Lebesgue spaces. *J. Math. Anal. Appl.*, 433 (2016), 176–199.

[4] Andersen N. and de Jeu M.: Real Paley-Wiener theorems and local spectral radius formulas. *Transactions of the American Mathematical Society,* 362(7), 3613-3640 (2010).

[5] Athreya S., Butkovsky O. and Mytnik L.: Strong existence and uniqueness for stable stochastic differential equations with distributional drift. *arXiv:1801.03473,* (2018).

[6] Bae J. and Kassmann M.: Schauder estimates in generalized Hölder spaces. *arXiv:1505.05498,* (2015).

[7] Bahouri H., Chemin J.-Y. and Danchin R.: Fourier Analysis and Nonlinear Partial Differential Equations. *Grundlehren der mathematischen Wissenschafen, Springer-Verlag,* Vol.343 (2011).

[8] Barles G., Chasseigne E. and Imbert C.: Hölder continuity of solutions of second-order non-linear elliptic integro-differential equations. *J. Eur. Math. Soc.*, 13 (2011), 1-26.

[9] Bass R. and Chen Z.: Stochastic differential equations for Dirichlet processes. *Probability Theory and Related Fields.* 121(3) (2001), 422-446.

[10] Bass R. and Chen Z.: Brownian motion with singular drift. *Annals of Prob.* 31(2) (2003), 791-817.

[11] Breiman L.: Probability. *Society for Industrial and Applied Mathematics. Classics in Applied Mathematics* (1992).

[12] Chen Z., Song, R. and Zhang X.: Stochastic flows for Lévy processes with Hölder drifts. To appear in *Revista Matematica Iberoamericana*, (2015).

[13] Chen Z. and Zhang X.: Uniqueness of stable-like processes. *arXiv preprint arXiv:1604.02681*, (2016).

[14] Chen Z., Zhang X. and Zhao G.: Well-posedness of supercritical SDE driven by Lévy processes with irregular drifts. *arXiv:1709.04632*, (2017).

[15] Debussche A. and Fournier N.: Existence of densities for stable-like driven SDE's with Hölder continuous coefficients. *J. Funct. Anal,* 264(8) (2013), 1757-1778.

[16] Dong H. and Kim D.: Schauder estimates for a class of non-local elliptic equations. *Discrete Contin. Dyn. Syst.* , 33 (2013) 2319-2347.

[17] Dong H., Jin T. and Zhang H.: Dini and Schauder estimates for nonlocal fully nonlinear parabolic equations with drifts. *Analysis and PDE* , 11.6(2018) 1487-1534.

[18] Engelbert H. and Schmidt W.: Strong Markov continuous local martingales and solutions of one-dimensional stochastic differential equations, I, II, III. *Math. Nachr.,* 143, 167-184; 144, 241-281; 151, 149-197 (1989, 1991).

[19] Elworhty K. and Li X.: Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* 125 (1) (1994), 252-286.

[20] Ethier S. and Kurtz T.: Markov processes: characterization and convergence. *John Wiley and Sons*, 282 (2009).

[21] Fedrizzi E. and Flandoli F.: Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics*, 83(3) (2011), 241-257.

[22] Fedrizzi E., Flandoli F., Priola E. and Vovelle J.: Regularity of stochastic Kinetic equations. *Electron. J. Probab.,* 22 (2017), 1–42.

[23] Figalli A.: Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *Journal of Functional Analysis, 254(1)* (2008), 109-153.

[24] Folland G.: Real Analysis: Modern Techniques and Their Applications. *A Wiley-Interscience Publication JOHN WILEY and SONS, INC.* (1999).

[25] Föllmer H.: Dirichlet processes. *Stochastic integrals,* 851 (1981), 476-478.

[26] Fourier N.: Finiteness of entropy for homogeneous Boltzmann equation with measure initial condition, *arXiv: 1203.0130.*

[27] Fourier N. and Printems, J.: Absolute continuity of some one-dimensional processes, *Bernoulli.* 16 (2010), 343-360.

[28] Flandoli F.: Random Perturbation of PDEs and Fluid Dynamic Models. *Lecture Notes in Mathematics. Springer, Heidelberg* (2011).

[29] Flandoli F., Gubinelli M. and Priola E.: Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.* 180 (2010), 1-53.

[30] Gyongy I. and Krylov N. V.: On the rate of convergence of splitting-up approximations for SPDEs. *Progress in Probability. Birkhauser Verlag, Basel*, 56 (2003), 301-321.

[31] Haadem S. and Proske F.: On the construction and Malliavin differentiability of solutions of Lévy noise driven SDE's with singular coefficients. *J. Fun. Anal.*, 266 (2014), 5321-5359.

[32] Ho K.: Strong maximal operator on mixed-norm spaces, *Ann Univ Ferrara.*, 62 (2016), 275-291

[33] Ikeda N. and Watanabe S.: Stochastic Differential Equations and Diffusion Processes. *North-Holland Publishing Company, Amsterdam Oxford New York* (1981).

[34] Krylov N. V.: On estimates of the maximum of a solution of a parabolic equation and estimates of the distribution of a semimartingale. *Math. USSR Sbornik*, 58 (1987).

[35] Krylov N. V.: A parabolic Littlewood-Paley inequality with applications to parabolic equations. *Topol. Methods Nonlinear Anal.*, 4 (1994), 355–364.

[36] Krylov N. V.: The heat equation in $L^q((0,T), L^p)$-spaces with weights. *SIAM J. Math. Anal.*, 32 (2001), 1117-1141.

[37] Krylov N. V. and Röckner M.: Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131 (2) (2005), 154–196.

[38] Kim K.: $L_q(L_p)$-theory of parabolic PDEs with variable coefficients. *Bull. Korean Math. Soc.,* 45 (2008), 169-190.

[39] Kurtz T.: Equivalence of stochastic equations and martingale problems. *Stochastic Analysis* (2010), 113-130.

[40] Lady$\check{z}$enskaja O., Solonnikov V. and Uralćeva N.: Linear and Quasi-linear Equations of Parabolic Type. *American Mathematical Society.* (1968).

[41] Lee H. and Trutnau G.: Existence and uniquness results for Itô's-SDEs with locally integrable drifts and Sobolev diffusion coefficients. *https://arxiv.org/abs/1708.01152*

[42] Lepeltier J. and Marchal B. Problème des martingales et équations différentielles stochastiques associées à un opérateur intégrodifférentiel. *Annales de l'IHP Probabilités et statistiques* 12(1) (1976), 43-103.

[43] Ling C., Michael Röckner and Zhu X.: SDEs with singular drifts and multiplicative noise on general space-time domains. *https://arxiv.org/pdf/1910.03989.pdf.*

[44] Ling C. and Xie. L.: Strong well-posedness for stochastic differential equations with coefficients in mixed-norm spaces. Preprint.

[45] Ling C. and Zhao G.: Nonlocal elliptic equation in Hölder space and the martingale problem. *https://arxiv.org/pdf/1907.00588v1.pdf*

[46] Liu W. and Röckner M.: Stochastic Partial Differential Equations: An Introduction.*Springer. Universitext.* (2010).

[47] Lipster R. and Shiryaev A.: Statistics of Random Processes. *"Nauka", Moscow, 1974 in Russian, English translation: Springer-Verlag, Berlin and New York,* (1977)

[48] Mikulevicius R. and Pragarauskas H.: On the Cauchy problem for certain integro-differential operators in Sobolev and Hölder spaces. *Lithuanian Math. Journal,* 32(2) (1992).

[49] Menoukeu P., Meyer B., Nilssen T., Proske F. and Zhang T.: A variational approach to the construction and Malliavin differentiability of strong solutions of SDEs. *Math. Ann.,* 357 (2013), 761-799.

[50] ∅ksendal B.: Stochastic Differential Equations. An Introduction with Applications. *Springer-Verlag Heidelberg New York* , (2000).

[51] Portenko N.: Generalized Diffusion Processes. *"Nauka", Moscow, 1982 in Russian, English translation: Amer. Math. Soc. Provdence, Rhode Island,* (1990).

[52] Priola E.: Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka Journal of Mathematics*, 49 (2012), 421-447.

[53] Priola E.: Stochastic flow for SDEs with jumps and irregular drift term. *Banach Center Publications 105(1)* (2015), 193-210.

[54] Protter P.: Stochastic Interation and Differentail Equations. *Stochatic Molldelling and Applied Probability, 21. Springer-Verlag, Berlin* (2005).

[55] Riahi L.: Comparison of Green functions and harmonic measures for parabolic operators. *Potential Anal.* 23 (4), (2005), 381-402.

[56] Schilling R., Sztonyk P. and Wang J.: Coupling property and gradient estimates of Lévy processes via the symbol. *Bernoulli,* , 18(4) (2012) 1128-1149.

[57] Silvestre L.: On the differentiability of the solution to an equation with drift and fractional diffusion. *Indiana Univ. Math. J.* , 61 (2012), 557-584.

[58] Song R. and Xie L.: Weak and strong well-posedness of critical and supercritical SDEs with singular coefficients. *arXiv preprint arXiv:1806.09033*. (2018).

[59] Stroock D. and Varadhan S.: Multidimensional diffusion processes. *Springer*, (2007).

[60] Tanaka H., Tsuchiya M. and Watanabe S.: Perturbation of drift-type for Lévy processes. *J. Math. Kyoto Univ.*, 14 (1974), 73-92

[61] Trevisan D.: Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electronic Journal of Probability,* , 21 (2016).

[62] Veretennikov A.: On the strong solutions of stochastic differential equations. *Theory Probab. Appl.*, 24 (1979), 354-366

[63] Veretennikov A.: On strong solutions and explicit formulas for solutions of stochastic integral equations. *Math. Ussr Sb.*, 39 (1981), 387-403

[64] Von der Lühe K.: Pathwise uniqueness for stochastic differential equations with singular drift and nonconstant diffusion. *PhD thesis. Bielefeld University.* (2018).

[65] Wang F.: Integrability conditions for SDEs and semi-linear SPDEs. *Annals of Prob.*, 2017+.

[66] Wang H. and Zhang Z.: A frequency localized maximum principle applied to the 2D quasi-geostrophic equation. *Communications in Mathematical Physics,* 301 (2011), 105-129.

[67] Xia P., Xie L, Zhang X. and Zhao G.: $L^q(L^p)$-theory of stochastic differential equations.

[68] Xie L.: Exponential ergodicity for stochastic Langevin equation with partial dissipative drift. *Applied Math. L.*, 79 (2018), 34-42.

[69] Xie L. and Zhang X.: Ergodicity of stochastic differential equations with jumps and singular coefficients. Accepted by *Ann. Inst Henri Poincare-Pr.*, https://arxiv.org/pdf/1705.07402.pdf.

[70] Xie L. and Zhang X.: Sobolev differential flows of SDEs with local Sobolev and super-linear growth coefficients. *Ann. Probab.* 44 (6) (2016), 3661-3687.

[71] Yamada T. and Watanabe S.: On the uniqueness of solutions of Stochastic differential equations. *I, II. J. Math. Kyoto Univ.* 11 (1971), 155-167, 553-563

[72] Zhang X.: Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients. *Stochastic Process. Appl.* 115 (11) (2005) 1805-1818.

[73] Zhang X.: Stochastic homomorphism flows of SDE with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* , 16(38) (2011), 1096-1116.

[74] Zhang X.: Stochastic differential equations with Sobolev drifts and driven by α-stable processes. *Ann. Inst. H. Poincare Probab. Statist.,* 49 (2013), 915-1231.

[75] Zhang X.: Stochastic differential equations with Sobolev coefficients and applications. *Annals of Applied Prob.*, 26 (5) (2016), 2697-2732.

[76] Zhang X. and Zhao G.: Heat kernel and ergodicity of SDEs with distributional drifts. *arXiv:1710.10537,* (2017).

[77] Zhang X. and Zhao G.: Singular Brownian Diffusion Processes. *Communications in Mathematics and Statistics,* 1-49, (2018).

[78] Zhang, X. and Zhao, G.: Dirichlet problem for supercritical nonlocal operators. *arXiv:1809.05712,* (2018).

[79] Zvonkin A.: A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sbornik*, 93 (135) (1974), 129-149.